

**This Page Is Inserted by IFW Operations  
and is not a part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

---

**Shears, B verly**

---

Fr m:  
Sent:  
T :  
Subj ct:

Devi, Sarvamangala  
Friday, January 31, 2003 7:36 AM  
Shears, Beverly  
09/699,224

Good morning Beverly:

Would you please perform a sequence and an interference search for SEQ ID NO: 1 in application SN 09/699,224?

Thanks.

S. DEVI, Ph.D.  
AU 1645  
CM1-7E15

Point of Contact:  
Beverly Shears  
Technical Info. Specialist  
CM1 1E05 Tel: 308-4994

7E12

**THIS PAGE BLANK (USPTO)**

09/699224

FILE 'REGISTRY' ENTERED AT 10:43:48 ON 31 JAN 2003  
L1 1 S IPVLDENGLFAP/SQSP

L1 ANSWER 1 OF 1 REGISTRY COPYRIGHT 2003 ACS  
RN 338797-97-0 REGISTRY  
CN L-Proline, L-isoleucyl-L-prolyl-L-valyl-L-leucyl-L-.alpha.-aspartyl-  
L-.alpha.-glutamyl-L-asparaginyglycyl-L-leucyl-L-phenylalanyl-L-  
alanyl- (9CI) (CA INDEX NAME)  
OTHER NAMES:  
CN 20: PN: WO0132699 SEQID: 1 claimed protein  
SQL 12

SEQ 1 IPVLDENGLF AP

HITS AT: 1-12

REFERENCE 1: 134:352269

FILE 'HCAPLUS' ENTERED AT 10:44:25 ON 31 JAN 2003  
L2 1 S L1

L2 ANSWER 1 OF 1 HCAPLUS COPYRIGHT 2003 ACS  
ACCESSION NUMBER: 2001:338560 HCAPLUS  
DOCUMENT NUMBER: 134:352269  
TITLE: Peptide mimics of conserved gonococcal epitopes  
and methods and compositions using them  
INVENTOR(S): Rice, Peter A.; Ngampasutadol, Jutamas; Gulati,  
Sunita  
PATENT ASSIGNEE(S): USA  
SOURCE: PCT Int. Appl., 57 pp.  
CODEN: PIXXD2  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
FAMILY ACC. NUM. COUNT: 1  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001032692	A2	20010510	WO 2000-US29749	20001027
WO 2001032692	A3	20020307		

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH,  
CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,  
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ,  
PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ,  
UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ,  
TM  
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH,  
CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE,  
BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.: US 1999-162491P P 19991029

AB The present invention relates to peptide mimics of a conserved  
gonococcal epitope of Neisseria gonorrhoeae, which epitope is not  
found on human blood group antigens. This invention also relates to  
methods and compns. using such peptide mimics for the prophylaxis of  
gonorrheal infections.

IT 338797-97-0, Ipvldenglfap peptide+  
RL: BAC (Biological activity or effector, except adverse); BSU

**THIS PAGE BLANK (USPTO)**

---

09/699224

(Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
(antigenic peptide mimics of conserved gonococcal epitopes and methods and compns. using them)

FILE 'HOME' ENTERED AT 10:44:37 ON 31 JAN 2003

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**

---



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2003, 07:59:04 : Search time 35 Seconds  
(without alignments)  
45.686 Million cell updates/sec

Title: US-09-699-224A-1

Perfect score: 63

Sequence: 1 IPELDENGLFAP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: A.Geneseq\_101002.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
24: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	63	100.0	12	22	AAB97116	Peptide mimic #1 O
2	42.5	67.5	15	22	AAB97125	Conserved gonococcal
3	41	65.1	232	21	AAV74788	Neisseria meningit
4	41	65.1	232	21	AAV74789	Neisseria meningit
5	41	65.1	724	22	AAU57330	Propionibacterium
6	40	63.5	224	22	ABB61313	Drosophila melanog
7	40	63.5	986	22	ABB57810	Drosophila melanog
8	39	61.9	92	22	AAE07086	Human gene 8 encod
9	39	61.9	93	21	AA656366	Arabidopsis thalia
10	39	61.9	100	21	AA656365	Arabidopsis thalia

11	39	61.9	135	22	AAE07058	Human gene 8 encod
12	39	61.9	135	22	AAE07085	Human gene 8 encod
13	39	61.9	225	22	AAB61361	Drosophila melanog
14	39	61.9	316	22	AA92033	C glutamicum prote
15	39	61.9	676	21	AAB23269	Balanus amphitrite
16	39	61.9	2893	19	AAW98828	H. pylori GHPD 148
17	39	61.9	2893	19	AAW71556	Helicobacter polyo
18	38	60.3	51	22	AA48120	Consensus thyriglo
19	38	60.3	313	22	AA682332	S. epidermidis ope
20	38	60.3	328	23	ABP38384	Staphylococcus epi
21	38	60.3	515	21	AA631867	Arabidopsis thalia
22	38	60.3	582	21	AA631866	Arabidopsis thalia
23	38	60.3	620	21	AA631865	Arabidopsis thalia
24	38	60.3	620	23	ABP91842	Herbicidally activ
25	38	60.3	644	21	AA648152	Arabidopsis thalia
26	38	60.3	711	21	AA648151	Arabidopsis thalia
27	38	60.3	712	17	AA95548	Thermotestable DNA-1
28	38	60.3	749	21	AA648150	Arabidopsis thalia
29	38	60.3	1663	23	ABB05714	Human testis deriv
30	37.5	59.5	115	23	ABP29318	Streptococcus poly
31	37	58.7	12	23	ABP79576	ADP-41 tryptic pe
32	37	58.7	12	23	ABP79576	Human ADP tryptic
33	37	58.7	174	22	AAU19596	Human diagnostic a
34	37	58.7	220	23	ABP65332	Drosophila melanog
35	37	58.7	230	23	AAU91139	Human secreted pro
36	37	58.7	236	22	AA95218	Human protein sequ
37	37	58.7	242	22	AA93934	Human protein sequ
38	37	58.7	251	23	AAU91140	Human secreted pro
39	37	58.7	317	21	AAV79242	Mortierella alpina
40	37	58.7	317	22	AA631687	Amino acid sequenc
41	37	58.7	317	23	AAU87820	M. alpina elongase
42	37	58.7	322	22	AAU40413	Human polypeptide
43	37	58.7	322	22	AA662389	Rat tricarboxylate
44	37	58.7	322	23	ABP79572	Alzheimer's diseas
45	37	58.7	322	23	ABP68039	Human Alzheimer's

## ALIGNMENTS

RESULT 1	
AAB97116	
ID AAB97116 standard; peptide; 12 AA.	
XX AAB97116:	
AC	
XX	
DT 07-AUG-2001 (first entry)	
XX	
DE Peptide mimic #1 of conserved gonococcal mab 2C7 epitope.	
XX	
KW Peptide mimic; vaccine; gonococcal; epitope; Neisseria gonorrhoeae;	
KW infection; monoclonal antibody 2C7; mab 2C7.	
XX	
OS Synthetic.	
XX	
PN WO200132692-A2.	
XX	
PD 10-MAY-2001.	
XX	
PF 27-OCT-2000; 2000WO-US29749.	
XX	
PR 29-OCT-1999; 99US-0162491.	
XX	
PA (RICE/) RICE P A.	
PA (NGAM/) NGAMPASUTADOL J.	
PA (GULA/) GULATI S.	
PI	
XX	
DR Rice PA. Ngampasutadol J, Gulati S;	
XX	
WPI; 2001-343473/36.	
XX	
PT New peptide mimics of conserved gonococcal epitopes not present in	
PT human blood group antigens, useful for prophylaxis of Neisseria	

PT gonorrhoeae infections -  
 XX  
 PS Claim 16; Fig 2; 57pp; English.  
 CC  
 CC The invention relates to novel peptide mimics of conserved gonococcal  
 CC epitopes which are not present in human blood group antigens. The  
 CC peptide mimics are useful for immunising against *Neisseria gonorrhoeae*  
 CC infections. The present sequence is a peptide mimic which binds to  
 CC an epitope of the *Neisseria gonorrhoeae* monoclonal antibody 2C7. It  
 CC was synthesised by random peptide display and, following five rounds of  
 CC positive selection with mAb 2C7, it was identified as being able to bind  
 CC mAb 2C7 by western blotting.  
 CC  
 SQ Sequence 12 AA;  
 Query Match 100.0%; Score 63; DB 22; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 IPVLDENGLEFAP 12  
 |||||  
 1 IPVLDENGLEFAP 12  
 RESULT 2  
 AAB97125  
 ID AAB97125 standard; peptide; 15 AA.  
 AC AAB97125;  
 XX  
 DT 07-AUG-2001 (first entry)  
 XX  
 DE Conserved gonococcal mAb 2C7 epitope peptide mimic fusion peptide.  
 XX  
 KW Peptide mimic; vaccine; gonococcal; epitope; *Neisseria gonorrhoeae*;  
 KM infection; monoclonal antibody 2C7; mAb 2C7; fusion.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200132692-A2.  
 PD 10-MAY-2001.  
 XX  
 PF 27-OCT-2000; 2000WO-US29749.  
 XX  
 PR 29-OCT-1999; 99US-0162491.  
 XX  
 PA (RICE/) RICE P A.  
 (NGAM/) NGAMPASUTADOL J.  
 (GULA/) GULATI S.  
 Rice PA, Ngampasutadol J, Gulati S;  
 WPI: 2001-343473/36.  
 PT New peptide mimics of conserved gonococcal epitopes not present in  
 PT human blood group antigens, useful for prophylaxis of *Neisseria*  
 PT gonorrhoeae infections -  
 XX  
 PS Claim 23; Fig 2; 57pp; English.  
 CC  
 CC The invention relates to novel peptide mimics of conserved gonococcal  
 CC epitopes which are not present in human blood group antigens. The  
 CC peptide mimics are useful for immunising against *Neisseria gonorrhoeae*  
 CC infections. The present sequence is a peptide mimic which binds to  
 CC an epitope of the *Neisseria gonorrhoeae* monoclonal antibody 2C7.  
 CC Peptides were synthesised by random peptide display and were subjected  
 CC to five rounds of positive selection with mAb 2C7 to identify those able  
 CC to bind mAb 2C7. Two cysteine flanking regions were added to the N- and  
 CC C-terminal of a peptide mimic produced by this method to generate the  
 CC present sequence, which is also an effective peptide mimic of the mAb  
 CC 2C7 epitope.  
 CC  
 SQ

SQ Sequence 15 AA;  
 Query Match 67.5%; Score 42.5; DB 22; Length 15;  
 Best Local Similarity 83.3%; Pred. No. 0.38;  
 Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 1 IPVLDENGLEFAP 12  
 |||||  
 4 IPVL-ENGLEFAP 14  
 Db  
 RESULT 3  
 AAY74788  
 ID AAY74788 standard; protein; 232 AA;  
 AC AAY74788;  
 XX  
 DT 21-MAR-2000 (first entry)  
 XX  
 DE *Neisseria meningitidis* ORF 261 protein sequence SEQ ID NO:1050.  
 XX  
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;  
 KM antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
 KW antibacterial; gene therapy.  
 XX  
 OS *Neisseria meningitidis*.  
 XX  
 PN WO957280-A2.  
 PD 11-NOV-1999.  
 XX  
 PF 30-APR-1999; 99WO-US09346.  
 XX  
 PR 01-MAY-1998; 98US-0083758.  
 PR 31-JUL-1998; 98US-0094869.  
 PR 02-SEP-1998; 98US-0098994.  
 PR 02-SEP-1998; 98US-0099062.  
 PR 09-OCT-1998; 98US-0103749.  
 PR 09-OCT-1998; 98US-0103794.  
 PR 09-OCT-1998; 98US-0103796.  
 PR 25-FEB-1999; 99US-0121528.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.  
 XX  
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
 PI Petersen J, Pizzo M, Rappoli R, Ratti G, Scalato E, Scarselli M;  
 PI Tettelin H, Venter JC;  
 XX  
 DR WPI: 2000-062150/05.  
 DR N-PSDB; AA253550.  
 XX  
 PT Novel *Neisseria* polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics -  
 XX  
 PS Claim 2; Page 604; 1453pp; English.  
 CC  
 CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941  
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides  
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent  
 CC PCR primers used in the exemplification of the present invention. The  
 CC polypeptides, the polynucleotides, antibodies and compositions of  
 CC the invention can be used as vaccines, as diagnostic reagents, and as  
 CC immunogenic compositions. The polypeptides can be used in the  
 CC manufacture of medicaments for treating or preventing infection due to  
 CC *Neisseria* bacteria (e.g. meningitis and septicemia), to detect the  
 CC presence of *Neisseria* bacteria, or to raise antibodies. They may also  
 CC be used to screen for agonists or antagonists, which may themselves  
 CC have use as antibacterial agents. The polynucleotides of the invention  
 CC may also be used in gene therapy protocols.  
 CC  
 SQ Sequence 232 AA;

Query Match 65.1%; Score 41; DB 21; Length 232;  
 Best Local Similarity 64.3%; Pred. No. 19;  
 Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

OY 1 IPVLDENGL--FAP 12  
 :|||: ||| |||  
 DB 194 LPVLESNGLDVFP 207

## RESULT 4

AA74789  
 ID AAY74789 standard; Protein; 232 AA.

AC AAY74789;  
 XX

DT 21-MAR-2000 (first entry)

DE Neisseria meningitidis ORF 261 protein sequence SEQ ID NO:1052.

Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
 antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
 antibacterial; gene therapy.

OS Neisseria meningitidis.

PN MO9957280-A2.

PD 11-NOV-1999.

PF 30-APR-1999; 99MO-US09346.

PR 01-MAY-1998; 98US-0083758.

PR 31-JUL-1998; 98US-0094869.

PR 02-SEP-1998; 98US-0098894.

PR 02-SEP-1998; 98US-0099062.

PR 09-OCT-1998; 98US-0103749.

PR 09-OCT-1998; 98US-0103794.

PR 09-OCT-1998; 98US-0103796.

PR 25-FEB-1999; 99US-0121528.

PA (CHIR) CHIRON CORP.

PA (GENO) INST GENOMIC RES.

PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;

PI Petersen J, Pizzi M, Rappuoli R, Ratti G, Scalato E, Scarselli M;

PI Tettelin H, Venter JC;

PI WPI; 2000-062150/05.

PI N-PSDB; AA253551.

XX Novel Neisserial polypeptides predicted to be useful antigens for

XX vaccines and diagnostics

XX Claim 2: Page 605; 1453pp; English.

XX AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941

XX represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides

XX and polypeptides. AA254577 to AA254576 and AA254616 to AA25473 represent

XX PCR primers used in the exemplification of the present invention. The

XX polypeptides, the polynucleotides, antibodies and compositions of

XX the invention can be used as vaccines, as diagnostic reagents, and as

XX immunogenic compositions. The polypeptides can be used in the

XX manufacture of medicaments for treating or preventing infection due to

XX Neisserial bacteria (e.g. meningitis and septicemia), to detect the

XX presence of Neisseria bacteria, or to raise antibodies. They may also

XX be used to screen for agonists or antagonists, which may themselves

XX have use as antibacterial agents. The polynucleotides of the invention

XX may also be used in gene therapy protocols.

XX Sequence 232 AA;

XX Query Match 65.1%; Score 41; DB 21; Length 232;

XX Best Local Similarity 64.3%; Pred. No. 19;

Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

OY 1 IPVLDENGL--FAP 12  
 :|||: ||| |||  
 DB 194 LPVLESNGLDVFP 207

## RESULT 5

AAU57330  
 ID AAU57330 standard; Protein; 724 AA.

AC AAU57330;  
 XX

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #18226.

SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 uvelitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001MO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORIXA CORP.

PI Sheikh YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'malsonneuve J, Zhang Y, Jen S, Carter D;

PI WPI; 2001-616774/71.

PI N-PSDB; AAS59582.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

XX vaccinating against and diagnosing infections, especially useful for

XX treating acne vulgaris

XX Example 1; SEQ ID NO 18525; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

XX polypeptides. The proteins and their associated DNA sequences are used in

XX the treatment, prevention and diagnosis of medical conditions caused by

XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

XX pustulosis, hypertosis and osteomyelitis), uvelitis and endophthalmitis.

XX P. acnes is also involved in infections of bone, joints and the central

XX nervous system, however it is particularly involved in the inflammatory

XX lesions associated with acne vulgaris. A method for detecting the

XX presence or absence of P. acnes in a patient comprises contacting a

XX sample with a binding agent that binds to the proteins of the invention

XX and determining the amount of bound protein in the sample. The

XX polypeptides may be used as antigens in the production of antibodies

XX specific for P. acnes proteins. These antibodies can be used to

XX downregulate expression and activity of P. acnes polypeptides and

XX therefore treat P. acnes infections. The antibodies may also be used as

XX diagnostic agents for determining P. acnes presence, for example, by

XX enzyme linked immunosorbent assay (ELISA).

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 724 AA;

XX Query Match 65.1%; Score 41; DB 22; Length 724;

XX Best Local Similarity 58.3%; Pred. No. 72;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 IPVLIDNGLFAP 12  
 :||:|||||  
 XX  
 DB 309 VPMLDNGMTP 320

RESULT 6  
 ABB61313  
 ID ABB61313 standard; Protein; 224 AA.  
 XX

AC ABB61313;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide spq ID NO 10731.

XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL05416.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 interactions -

XX Disclosure; SEQ ID NO 10731; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABBS7737-ABBS72072).

CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 224 AA;

XX Query Match 63.5%; Score 40; DB 22; Length 224;  
 Best Local Similarity 60.0%; Pred. No. 27;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 IPVLIDNGLF 10  
 :||:|||||  
 XX  
 DB 57 VPMLDNGMTP 66

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide spq ID NO 222.

XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL01913.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 interactions -

XX Disclosure; SEQ ID NO 222; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABBS7737-ABBS72072).

CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 986 AA;

XX Query Match 63.5%; Score 40; DB 22; Length 986;  
 Best Local Similarity 70.0%; Pred. No. 16+02;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 IPVLIDNGLF 10  
 :||:|||||  
 XX  
 DB 297 IPILEDGTF 306

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX



PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140354.  
 PR 28-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 02-JUL-1999; 99US-0142154.  
 PR 06-JUL-1999; 99US-0142055.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147316.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154479.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158332.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.



PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148365.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
	07-SEP-1999;	99US-0152363.
	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Db	51	LP1EENGIFS	61
RESULT 11			
AAE07058			
ID	AAE07058	standard; Protein; 135 AA.	
AC	AAE07058;		
DT	16-OCT-2001	(first entry)	
XX			
DE	Human gene 8 encoded secreted protein HHTLB76, SEQ ID NO:75.		
XX			
KW	Human secreted protein; proliferative disorder; cancer; tumour;		
KW	focal abnormality; developmental abnormality; haematopoietic disorder;		
KW	immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;		
KW	inflammation; allergy; neurological disorder; Alzheimer's disease;		
KW	Parkinson's disease; cognitive disorder; schizophrenia; asthma;		
KW	skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;		
KW	cardiovascular disorder; angiotensin disorder; kidney disorder;		
KW	gastrointestinal disorder; pregnancy-related disorder;		
KW	endocrine disorder; infection; wound healing; vulnerability;		
KW	cell culture; chemotaxis; food additive; gene therapy;		
KW	blinding partner identification.		
XX			
OS	Homo sapiens.		
XX			
FT	Key	Location/Qualifiers	
FT	Peptide	1..24	
FT	Protein	/label= Signal_peptide	
FT		25..135	
XX		/label= Mature_human_secreted_protein	
XX	WO200154708-A1.		
PD	02-AUG-2001.		
XX			
XX	17-JAN-2001; 2001WO-US01434.		
XX			
PR	31-JAN-2000; 2000US-0179065.		
PR	04-FEB-2000; 2000US-0180628.		
PR	18-AUG-2000; 2000US-0226279.		
PR	05-DEC-2000; 2000US-0251988.		
PR	05-JAN-2001; 2001US-0259678.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Rosen CA, Komatsuoullis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;		
PI	Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M,		
PI	Ni J, Ruben SM, Barash SC;		
XX			
DR	WPI: 2001-488743/53.		
DR	N-PSDB: AAD13352.		
XX			
PT	New isolated nucleic acids and polypeptides, useful for diagnosing,		
PT	treating and/or preventing human diseases and disorders -		
XX			
PS	Claim 11; Page 494; 558pp; English.		
XX			
CC	AAID13345-AAID13401 represent cDNAs corresponding to 22 human secreted		
CC	protein genes, and AAE07051-AAE07105 represent the proteins they encode.		
CC	AAE07106-AAE07129 represent human secreted protein fragments or variants		
CC	the genes and their secreted proteins are useful for preventing,		
CC	treating or ameliorating medical conditions, e.g., by protein or gene		
CC	therapy. Pathological conditions can be diagnosed by determining the		
CC	amount of the new protein in a sample or by determining the presence of		
CC	mutations in the new genes. Specific uses are described for each of the		
CC	22 genes, based on the tissues in which they are most highly expressed,		
CC	and include developing products for the diagnosis or treatment of		
CC	proliferative disorders, cancer, tumours, foetal and developmental		
CC	abnormalities, haematopoietic disorders, diseases of the immune system,		
CC	AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,		
CC	allergies, neurological disorders (e.g., Alzheimer's disease,		



CC	Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC	skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC	cardiovascular disorders, angioedema, disorders, kidney disorders,
CC	gastrointestinal disorders, pregnancy-related disorders, endocrine
CC	disorders, and infections. The proteins can also be used to aid wound
CC	healing and epithelial cell proliferation, to prevent skin aging due to
CC	sunburn, to maintain organs before transplantation, for supporting cell
CC	culture of primary tissues, to regenerate tissues, to identify their
CC	cognate ligands or binding partners, and in chemotaxis, and can be used
CC	as a food additive or preservative to modify storage properties.
CC	Antibodies specific for a protein of the invention can be used in
CC	alleviating symptoms associated with the disorders mentioned above, and
CC	in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC	immunosorbent assay (ELISA). The present sequence represents a human
CC	secreted protein of the invention.
XX	
SQ	Sequence 135 AA;
Query Match	61.9%; Score 39; DB 22; Length 135;
Best Local Similarity	63.6%; Pred. No. 23;
Matches 7; Conservative	1; Mismatches 3; Indels 0; Gaps 0;
OY	2 PVLIDNGLFAP 12
	1:1111111
Dd	46 PVLQETGAFAF 56
RESULT 12	
ID	AAE07085
AC	AAE07085 standard; Protein; 135 AA.
XX	
DT	16-OCT-2001 (first entry)
XX	
DE	Human gene 8 encoded secreted protein HHTLB76, SEQ ID NO:102.
XX	
KW	Human secreted protein; proliferative disorder; cancer; tumour;
KW	fetal abnormality; developmental abnormality; haematopoietic disorder;
KW	immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW	inflammation; allergy; neurological disorder; Alzheimer's disease;
KW	Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW	skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW	cardiovascular disorder; angioedema; kidney disorder;
KW	gastrointestinal disorder; pregnancy-related disorder;
KW	endocrine disorder; infection; wound healing; vulnery;
KW	cell culture; chemotaxis; food additive; gene therapy;
KW	binding partner identification.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	Protein
FT	Misc-difference
FT	125
FT	/label= Unknown
FT	/note= "Encoded by GYG"
XX	
PN	WO200154708-A1.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US01434.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	18-AUG-2000; 2000US-0226279.
PR	05-DEC-2000; 2000US-0251988.
PR	05-JAN-2001; 2001US-0259678.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.

Pt	Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS,
Pf	Moore PA, Wel P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M,
Pn	Ni J, Ruben SM, Barash SC;
Dl	MPI; 2001-488743/53.
Dr	N-PsDB; AAD13379.
Px	New isolated nucleic acids and polypeptides, useful for diagnosing,
Pt	treating and/or preventing human diseases and disorders -
Px	Claim 11; Page 521; 558pp; English.
Pd	AAD13345-AAD13401 represent cDNAs corresponding to 22 human secreted
Cc	protein genes, and AAEO7051-AAE07105 represent the proteins they encode.
Cc	AAEO7106-AAEO7129 represent human secreted protein fragments or variants.
Cc	The genes and their secreted proteins are useful for preventing,
Cc	treatng or ameliorating medical conditions, e.g., by protein or gene
Cc	therapy. Pathological conditions can be diagnosed by determining the
Cc	amount of the new protein in a sample or by determining the presence of
Cc	mutations in the new genes. Specific uses are described for each of the
Cc	22 genes, based on the tissues in which they are most highly expressed,
Cc	and include developing products for the diagnosis or treatment of
Cc	proliferative disorders, cancer, tumours, foetal and developmental
Cc	abnormalities, haematopoietic disorders, diseases of the immune system,
Cc	AIDS, autoimmune diseases (e.g., Rheumatoid arthritis), inflammation,
Cc	allergies, neurological disorders (e.g., Alzheimer's disease,
Cc	Parkinson's disease), cognitive disorders, schizophrenia, asthma,
Cc	skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
Cc	cardiovascular disorders, angiotonic disorders, kidney disorders,
Cc	gastrointestinal disorders, pregnancy-related disorders, endocrine
Cc	disorders, and infections. The proteins can also be used to aid wound
Cc	healing and epithelial cell proliferation, to prevent skin aging due to
Cc	sunburn, to maintain organs before transplantation, for supporting cell
Cc	culture of primary tissues, to regenerate tissues, to identify their
Cc	cognate ligands or binding partners, and in chemotaxis, and can be used
Cc	as a food additive or preservative to modify storage properties.
Cc	Antibodies specific for a protein of the invention can be used in
Cc	alleivating symptoms associated with the disorders mentioned above, and
Cc	in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
Cc	immunoscort assay (ELISA). The present sequence represents a human
Sq	secreted protein of the invention.
Xx	Sequence    135 AA:
Qy	Query Match                  61.9%; Score 39; DB 22; Length 135;
Bt	Best Local Similarity         63.6%; Pred. No. 23;
Mtches	Matches      7; Conservative      1; Mismatches    3; Indels      0; Gaps      0
OY	2 PYLDNGLEFAP 12   :       Dl 46 PLLOETGAFAFP 56
RESULT 13	
ID	ABB61361 standard; Protein; 225 AA.
AC	ABB61361;
DT	26-MAR-2002 (first entry)
DE	Drosophila melanogaster polypeptide SEQ ID NO 10875.
KW	Drosophilal development biology; cell signalling; insecticide; Pharmaceutical.
OS	Drosophila melanogaster.
PN	WO200171042-A2.
PD	27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.  
XX  
XX 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI WPI: 2001-656860/75.  
DR N-PSDB; ABL05464.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from *Drosophila* and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Disclosure; SEQ ID NO 10875; 21bp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA  
CC sequences (ABU1840-ABU16175) and the encoded proteins  
CC (ABU57737-ABU72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 225 AA;  
SQ  
Query Match 61.9%; Score 39; DB 22; Length 225;  
Best Local Similarity 87.5%; Pred. No. 42;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IPYLDENG 8  
DB 56 IPYLDENG 63  
RESULT 14  
AAG92033  
ID AAG92033 standard; Protein; 316 AA.  
XX  
XX AAG92033;  
AC  
XX  
XX 26-SEP-2001 (first entry)  
DT  
XX C glutamicum protein fragment SEQ ID NO: 5787.  
KM  
KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
XX organic acid synthesis.  
XX  
XX Corynebacterium glutamicum.  
OS  
XX  
XX EP1108790-A2.  
PN  
XX  
XX 20-JUN-2001.  
PD  
XX  
XX 18-DEC-2000; 2000EP-0127688.  
PF  
XX  
XX 16-DEC-1999; 99JP-0377484.  
PR 07-APR-2000; 2000JP-0159162.  
PR 03-AUG-2000; 2000JP-0280988.  
XX  
XX  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX  
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX  
XX WPI: 2001-376931/40.  
DR N-PSDB; AAH67252.

XX  
XX Novel polynucleotides derived from Coryneform bacteria, for identifying  
PT mutation point of a gene, measuring expression of a gene, analysing  
PT expression profile or pattern of a gene and identifying homologous gene  
PT  
XX  
XX Claim 17; SEQ ID NO: 5787; 246bp + Sequence Listing; English.  
PS  
XX  
XX The present invention provides a number of nucleotide and protein  
CC sequences from the Coryneform bacterium *Corynebacterium glutamicum*. These  
CC are useful for identifying the mutation point of a gene derived from a  
CC mutant of coryneform bacterium, measuring expression amount and  
CC analysing the expression profile or expression pattern of a gene derived  
CC from Coryneform bacterium, and identifying a homologue of a gene derived  
CC from coryneform bacterium. Coryneform bacteria are useful for producing  
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,  
CC particularly L-lysine. The present sequence is a protein described  
CC in the exemplification of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC European Patent Office.  
XX  
XX Sequence 316 AA;  
SQ  
Query Match 61.9%; Score 39; DB 22; Length 316;  
Best Local Similarity 77.8%; Pred. No. 63;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 4 LDENGKFP 12  
DB 167 VDENGKFP 175  
RESULT 15  
AAB23269  
ID AAB23269 standard; Protein; 676 AA.  
XX  
XX AAB23269;  
AC  
XX  
XX 02-FEB-2001 (first entry)  
DT  
XX  
XX Balanus amphitrite adhesion/metamorphosis-related protein Bcs-3.  
DE  
XX  
XX Adhesion/metamorphosis-related protein Bcs-3; barnacle; larva-specific;  
KM adhesion inhibition; metamorphosis inhibition; compound screening;  
KM antifouling composition.  
XX  
XX Balanus amphitrite.  
OS  
XX  
XX JP2000228985-A.  
PN  
XX  
XX 22-AUG-2000.  
PD  
XX  
XX 09-FEB-1999; 99JP-0031067.  
PF  
XX  
XX 09-FEB-1999; 99JP-0031067.  
PR  
XX  
XX (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.  
PA  
XX  
XX WPI: 2000-649634/63.  
DR N-PSDB; AAA97603.  
DR  
XX  
XX Novel barnacle kiris larva-specific adhesion/metamorphosis-related  
PT gene useful for identifying inhibitors of barnacle adhesion -  
PT  
XX  
XX Claim 3; Page 15-17; 32pp; Japanese.  
PS  
XX  
XX The invention relates to six larva-specific adhesion/metamorphosis-  
CC related genes from the barnacle *Balanus amphitrite* (CDNAs given in  
CC AA97601-A97606) and to the proteins they encode (AAB23267-B23272).  
CC The genes and the proteins can be used for screening for a  
CC substance that inhibits the adhesion or metamorphosis of barnacle  
CC larvae, which may be useful in antifouling compositions for use in

CC the shipping industry. The present sequence represents the Balanus  
 CC amphitrite adhesion/metamorphosis-related protein Bcs-3.  
 XX

SQ Sequence 676 AA;

Query Match 61.9%; Score 39; DB 21; Length 676;  
 Best Local Similarity 87.5%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 DENGFLAP 12  
 |||||

Db 563 DENGFLAP 570

Search completed: January 31, 2003, 07:59:49  
 Job time : 36 secs

**THIS PAGE BLANK (USPTO)**

---

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2003, 07:59:05 : Search time 14 Seconds  
(without alignments)  
25.220 Million cell updates/sec

Title: US-09-699-224A-1  
Perfect score: 63  
Sequence: 1 IPVLIDENGIFAP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues  
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: Issued Patents\_AA:\*  
2: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/backfillsl.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Query length	ID	Description
1	38	60.3	322	US-09-359-161-7	Sequence 7, Appl1
2	38	60.3	328	US-09-134-001C-3229	Sequence 3229, Ap
3	38	60.3	712	US-09-708-426-9	Sequence 9, Appl1
4	37	58.7	104	US-08-946-329A-97	Sequence 97, Appl
5	37	58.7	178	US-09-145-828A-24	Sequence 24, Appl
6	37	58.7	280	US-09-145-828A-21	Sequence 21, Appl
7	37	58.7	289	US-09-145-828A-17	Sequence 17, Appl
8	37	58.7	293	US-09-145-828A-12	Sequence 12, Appl
9	37	58.7	317	US-09-145-828A-7	Sequence 7, Appl1
10	37	58.7	318	US-09-145-828A-19	Sequence 19, Appl
11	36	57.1	7257	US-09-335-409-5	Sequence 5, Appl1
12	36	57.1	7257	US-09-568-102-5	Sequence 5, Appl1
13	36	57.1	7257	US-09-567-969-5	Sequence 5, Appl1
14	36	57.1	7257	US-09-568-480-5	Sequence 5, Appl1
15	36	57.1	7257	US-09-568-480-5	Sequence 5, Appl1
16	36	57.1	7257	US-09-568-480-5	Sequence 5, Appl1
17	36	57.1	7257	US-09-568-480-5	Sequence 5, Appl1
18	35	55.6	274	US-09-567-899-5	Sequence 5, Appl1
19	35	55.6	274	US-08-177-109A-58	Sequence 58, Appl
20	35	55.6	460	US-08-687-706-58	Sequence 58, Appl
21	34.5	54.8	1042	US-09-647-540A-2	Sequence 2, Appl1
22	34.5	54.8	1043	US-08-928-361B-11	Sequence 11, Appl
23	34.5	54.8	1721	US-08-928-361B-30	Sequence 30, Appl
24	34.5	54.8	1721	US-08-700-651-5	Sequence 5, Appl1
25	34.5	54.8	1837	US-08-928-361B-5	Sequence 5, Appl1
26	34	54.0	88	US-08-858-207A-504	Sequence 504, Ap
27	34	54.0	178	US-08-938-669A-29	Sequence 29, Appl

28	34	54.0	437	2	US-08-737-716-2	Sequence 2, Appl1
29	34	54.0	457	2	US-08-847-900-3	Sequence 3, Appl1
30	34	54.0	498	2	US-08-858-207A-310	Sequence 310, App
31	34	54.0	550	2	US-08-844-058-2	Sequence 2, Appl1
32	34	54.0	880	2	US-08-483-101-14	Sequence 14, Appl
33	33	52.4	28	1	US-08-299-249A-12	Sequence 12, Appl
34	33	52.4	237	6	5212074-7	Patent No. 5212074
35	33	52.4	246	1	US-08-286-888A-1	Sequence 1, Appl1
36	33	52.4	246	1	US-08-294-386C-1	Sequence 1, Appl1
37	33	52.4	246	1	US-08-284-386C-3	Sequence 3, Appl1
38	33	52.4	246	1	US-08-299-249A-1	Sequence 1, Appl1
39	33	52.4	246	1	US-08-299-249A-10	Sequence 10, Appl
40	33	52.4	246	1	US-08-590-708-1	Sequence 1, Appl1
41	33	52.4	246	5	PCT-US94-06543-1	Sequence 1, Appl1
42	33	52.4	246	5	PCT-US95-10224-1	Sequence 1, Appl1
43	33	52.4	246	5	PCT-US95-10224-3	Sequence 3, Appl1
44	33	52.4	263	4	US-09-134-001C-4638	Sequence 4638, Ap
45	33	52.4	311	4	US-09-282-305-16	Sequence 16, Appl

# ALIGNMENTS

```

RESULT 1
US-09-359-161-7
: Sequence 7, Application US/09359161A
: Patent No. 6342656
: GENERAL INFORMATION:
: APPLICANT: Bradford, Kent J.
: APPLICANT: Danah, Peetambar
: APPLICANT: Yang, Hong
: APPLICANT: Cooley, Michael
: APPLICANT: Downie, Bruce
: APPLICANT: Gee, Oliver
: TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses
: TITLE OF INVENTION: to Stress Conditions in Plants
: FILE REFERENCE: 023070-095900US
: CURRENT APPLICATION NUMBER: US/09/359,161A
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: Patent Ver. 2.1
: SEQ ID NO 7
: LENGTH: 322
: TYPE: PRT
: ORGANISM: Saccharomyces cerevisiae
: FEATURE:
: OTHER INFORMATION: yeast sucrose nonfermenting protein kinase 1
: OTHER INFORMATION: kinase subunit (SNF1)
US-09-359-161-7

Query Match      60.3%   Score 38;   DB 4;   Length 322;
Best Local Similarity 62.5%   Pred. No. 23;
Matches 5;   Conservative 3;   Mismatches 0;   Indels 0;   Caps 0;

Oy      1 IPVLIDENG 8
Db      223 VPIIDENG 230

RESULT 2
US-09-134-001C-3229
: Sequence 3229, Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: CURRENT FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08

```

;; PRIOR APPLICATION NUMBER: US 60/055,779  
;; PRIOR FILING DATE: 1997-08-14  
;; NUMBER OF SEQ ID NOS: 5674  
;; SEQ ID NO 3229  
;; LENGTH: 328  
;; TYPE: PRT  
;; ORGANISM: *Staphylococcus epidermidis*  
US-09-134-001C-3229

Query Match      60.3%; Score 38; DB 4; Length 328;  
Best Local Similarity 77.8%; Pred. No. 23;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 PVIDENGLF 10  
      1:|||||111  
Db      79 PVIDENGLF 87

RESULT 3  
US-09-708-426-9  
Sequence 9, Application US/09708426  
Patent No. 6444429

;; GENERAL INFORMATION:  
;; APPLICANT: HAN, YE-SUN  
;; APPLICANT: YU, YEON-GYU  
;; APPLICANT: LIM, JAE-HWAN  
;; TITLE OF INVENTION: GENE CODING FOR DNA LIGASE OF HYPERTHERMOPHILIC BACTERIA AQUIFEX  
;; TITLE OF INVENTION: PYROPHILUS AND PROTEIN EXPRESSED THEREFROM  
;; FILE REFERENCE: 199579050  
;; CURRENT APPLICATION NUMBER: US/09/708,426  
;; CURRENT FILING DATE: 2000-11-09  
;; PRIOR APPLICATION NUMBER: KR99-49591  
;; PRIOR FILING DATE: 1999-11-10  
;; NUMBER OF SEQ ID NOS: 12  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 9  
;; LENGTH: 712  
;; TYPE: PRT  
;; ORGANISM: *Rhodothermus marinus*  
US-09-708-426-9

Query Match      60.3%; Score 38; DB 4; Length 712;  
Best Local Similarity 77.8%; Pred. No. 58;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 IPVLDENGL 9  
      1:|||||111  
Db      697 IPVLDENGL 705

SUPT 4  
US-08-946-329A-97  
Sequence 97, Application US/08946329A  
Patent No. 6057091

;; GENERAL INFORMATION:  
;; APPLICANT: Beachy, Philip A.  
;; APPLICANT: Porter, Jeffrey A.  
;; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
;; NUMBER OF SEQUENCES: 109  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Fish & Richardson P.C.  
;; STREET: 4225 Executive Square, Suite 1400  
;; CITY: La Jolla  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 92037  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: Windows 95  
;; SOFTWARE: FastSeq for Windows Version 2.0b  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/946,329A

;; FILING DATE: 07-OCT-1997  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 60/061,323  
;; FILING DATE: 07-OCT-1996  
;; APPLICATION NUMBER: 08/729,743  
;; FILING DATE: 10-JUL-1996  
;; APPLICATION NUMBER: 08/567,357  
;; FILING DATE: 04-DEC-1995  
;; APPLICATION NUMBER: 08/349,498  
;; FILING DATE: 02-DEC-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Hallie, Lisa A.  
;; REGISTRATION NUMBER: 38,347  
;; REFERENCE/DOCKET NUMBER: 07265/140001  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 619/678-5070  
;; TELEFAX: 619/678-5099  
;; INFORMATION FOR SEQ ID NO: 97:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 104 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-946-329A-97

Query Match      58.7%; Score 37; DB 3; Length 104;  
Best Local Similarity 72.7%; Pred. No. 9.2;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 IPVLDENGLFA 11  
      1:|||||11  
Db      40 IMVLDENGLVA 50

RESULT 5  
US-09-145-828A-24  
Sequence 24, Application US/09145828A  
Patent No. 6403349

;; GENERAL INFORMATION:  
;; APPLICANT: Abbott Laboratories  
;; APPLICANT: Mukerji, Pradip  
;; APPLICANT: Leonard, Amanda E. Y.  
;; APPLICANT: Huang, Yung-Sheng  
;; APPLICANT: Thurmond, Jennifer  
;; APPLICANT: Kirchner, Stephen J.  
;; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF  
;; FILE REFERENCE: 6407.US.01  
;; CURRENT APPLICATION NUMBER: US/09/145,828A  
;; CURRENT FILING DATE: 1998-09-02  
;; NUMBER OF SEQ ID NOS: 30  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 24  
;; LENGTH: 178  
;; TYPE: PRT  
;; ORGANISM: *Saccharomyces cerevisiae*  
US-09-145-828A-24

Query Match      58.7%; Score 37; DB 4; Length 178;  
Best Local Similarity 60.0%; Pred. No. 17;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 IPVLDENGLF 10  
      1:|||||1111  
Db      3 VPILARNGLF 12

RESULT 6  
US-09-145-828A-21  
Sequence 21, Application US/09145828A  
Patent No. 6403349

;; GENERAL INFORMATION:  
;; APPLICANT: Abbott Laboratories

```
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Thurmond, Jennifer
; APPLICANT: Kirchner, Stephen J.
; APPLICANT: Parker-Barnes, Jennifer M.
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
; FILE REFERENCE: 6407.US.01
; CURRENT APPLICATION NUMBER: US/09/145,828A
; CURRENT FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 280
; TYPE: PRF
; ORGANISM: Mortierella alpina
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (280)...(280)
; OTHER INFORMATION: Xaa = Unknown or other at position 280
US-09-145-828a-21
```

```
Query Match          58.7%; Score 37; DB 4; Length 280;
Best Local Similarity 60.0%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 IPVDENGLEF 10
    :|:| | | | |
Db 70 VPILARNGLF 79
```

```
RESULT 7
US-09-145-828a-17
; Sequence 17, Application US/09145828A
; Patent No. 6403349
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Thurmond, Jennifer
; APPLICANT: Kirchner, Stephen J.
; APPLICANT: Parker-Barnes, Jennifer M.
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
; FILE REFERENCE: 6407.US.01
; CURRENT APPLICATION NUMBER: US/09/145,828A
; CURRENT FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: PRF
; ORGANISM: Mortierella alpina
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (289)...(289)
; OTHER INFORMATION: Xaa = Unknown or other at position 289
US-09-145-828a-17
```

```
Query Match          58.7%; Score 37; DB 4; Length 289;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 IPVDENGLEF 10
    :|:| | | | |
Db 79 VPILARNGLF 88
```

```
RESULT 8
US-09-145-828a-12
; Sequence 12, Application US/09145828A
; Patent No. 6403349
; GENERAL INFORMATION:
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
```

```
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Thurmond, Jennifer
; APPLICANT: Kirchner, Stephen J.
; APPLICANT: Parker-Barnes, Jennifer M.
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
; FILE REFERENCE: 6407.US.01
; CURRENT APPLICATION NUMBER: US/09/145,828A
; CURRENT FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 293
; TYPE: PRF
; ORGANISM: Mortierella alpina
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (293)...(293)
; OTHER INFORMATION: Xaa = Unknown or other at position 293
US-09-145-828a-12
```

```
Query Match          58.7%; Score 37; DB 4; Length 293;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 IPVDENGLEF 10
    :|:| | | | |
Db 83 VPILARNGLF 92
```

```
RESULT 9
US-09-145-828a-7
; Sequence 7, Application US/09145828A
; Patent No. 6403349
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Thurmond, Jennifer
; APPLICANT: Kirchner, Stephen J.
; APPLICANT: Parker-Barnes, Jennifer M.
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
; FILE REFERENCE: 6407.US.01
; CURRENT APPLICATION NUMBER: US/09/145,828A
; CURRENT FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 317
; TYPE: PRF
; ORGANISM: Mortierella alpina
US-09-145-828a-7
```

```
Query Match          58.7%; Score 37; DB 4; Length 317;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 IPVDENGLEF 10
    :|:| | | | |
Db 108 VPILARNGLF 117
```

```
RESULT 10
US-09-145-828a-19
; Sequence 19, Application US/09145828A
; Patent No. 6403349
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
```

```
APPLICANT: Huang, Yung-Sheng
APPLICANT: Thurmond, Jennifer
APPLICANT: Kirchner, Stephen J.
APPLICANT: Parker-Barnes, Jennifer M.
TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
FILE REFERENCE: 6407.US.01
CURRENT APPLICATION NUMBER: US/09/145,828A
CURRENT FILING DATE: 1998-09-02
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19
LENGTH: 318
TYPE: PRT
ORGANISM: Mortierella alpina
FEATURE:
NAME/KEY: VARIANT
LOCATION: (318)...(318)
OTHER INFORMATION: Xaa - Unknown or other at position 318
US-09-145-828A-19
```

```
Query Match          58.7%; Score 37; DB 4; Length 318;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 LPVIDENGLF 10
    :|:| | | | |
Db 108 VPILARNGLF 117
```

```
RESULT 11
US-09-335-409-5
; Sequence 5, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/335,409
CURRENT FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 7257
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-335-409-5
```

```
Query Match          57.1%; Score 36; DB 3; Length 7257;
Best Local Similarity 87.5%; Pred. No. 2.1e+03;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 4 LDENGLFA 11
    |||:| | | |
Db 7238 LDDEGLFA 7245
```

```
RESULT 12
US-09-568-102-5
; Sequence 5, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
```

```
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,102
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 7257
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-568-102-5
```

```
Query Match          57.1%; Score 36; DB 4; Length 7257;
Best Local Similarity 87.5%; Pred. No. 2.1e+03;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 4 LDENGLFA 11
    |||:| | | |
Db 7238 LDDEGLFA 7245
```

```
RESULT 13
US-09-567-969-5
; Sequence 5, Application US/09567969
; Patent No. 6355457
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/567,969
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 7257
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-567-969-5
```

```
Query Match          57.1%; Score 36; DB 4; Length 7257;
Best Local Similarity 87.5%; Pred. No. 2.1e+03;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 4 LDENGLFA 11
    |||:| | | |
Db 7238 LDDEGLFA 7245
```

```
RESULT 14
US-09-568-480-5
; Sequence 5, Application US/09568480
; Patent No. 6355458
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,480
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
```



```

; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-480-5

```

```

Query Match          57.1%; Score 36; DB 4; Length 7257;
Best Local Similarity 87.5%; Pred. NO. 2.1e+03;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 LDENGLFA 11
      111:1111
Db      7238 LDEDGLFA 7245

```

# RESULT 15

```

US-09-568-486-5
Sequence 5, Application US/09568486
Patent No. 6355459

```

```

; GENERAL INFORMATION:
; APPLICANT: Schnupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHIIONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,486
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-486-5

```

```

Query Match          57.1%; Score 36; DB 4; Length 7257;
Best Local Similarity 87.5%; Pred. NO. 2.1e+03;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 LDENGLFA 11
      111:1111
Db      7238 LDEDGLFA 7245

```

Search completed: January 31, 2003, 08:00:28  
Job time : 15 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2003, 08:00:34 ; Search time 11 Seconds  
(without alignments)  
22.013 Million cell updates/sec

Title: US-09-699-224A-1  
Perfect score: 63  
Sequence: 1 IPIVDENGLEFAP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues  
Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	61.9	316	9	US-09-738-626-5787
2	37	58.7	174	10	US-09-903-456-40
3	37	58.7	178	10	US-09-903-456-26
4	37	58.7	219	10	US-09-903-456-38
5	37	58.7	280	10	US-09-903-456-28
6	37	58.7	286	10	US-09-903-456-59
7	37	58.7	289	10	US-09-903-456-21
8	37	58.7	289	10	US-09-903-456-34
9	37	58.7	291	10	US-09-903-456-36
10	37	58.7	293	10	US-09-903-456-19
11	37	58.7	301	10	US-09-903-456-33
12	37	58.7	317	10	US-09-903-456-13
13	37	58.7	318	10	US-09-903-456-25
14	37	58.7	322	9	US-09-990-415A-8
15	37	58.7	436	9	US-09-978-295A-442
16	37	58.7	436	9	US-09-978-697-442
17	37	58.7	436	9	US-09-978-192A-442
18	37	58.7	436	9	US-09-999-832A-442
19	37	58.7	436	9	US-09-978-189-442

20	37	58.7	436	9	US-10-028-072-404	Sequence 404, App
21	37	58.7	436	10	US-09-745-763-138	Sequence 138, App
22	37	58.7	926	9	US-10-023-437-57	Sequence 57, App1
23	36	57.1	196	9	US-09-738-626-6624	Sequence 6624, Ap
24	36	57.1	313	9	US-10-174-590-554	Sequence 554, App
25	36	57.1	313	9	US-10-176-758-554	Sequence 554, App
26	36	57.1	313	9	US-10-175-737-554	Sequence 554, App
27	36	57.1	313	12	US-10-052-586-554	Sequence 554, App
28	36	57.1	429	9	US-09-738-626-5318	Sequence 5318, Ap
29	36	57.1	563	10	US-09-789-561-83	Sequence 83, App1
30	36	57.1	7257	9	US-10-014-717-5	Sequence 5, App1
31	35	55.6	111	9	US-10-042-894A-21	Sequence 21, App1
32	35	55.6	193	10	US-09-880-192-56	Sequence 56, App1
33	35	55.6	424	10	US-09-815-242-12587	Sequence 12587, A
34	35	55.6	425	10	US-09-815-242-12587	Sequence 5485, Ap
35	35	55.6	458	9	US-10-008-016-4	Sequence 4, App1
36	35	55.6	691	9	US-09-712-363-263	Sequence 263, App
37	35	55.6	906	10	US-09-815-242-10567	Sequence 10567, A
38	35	55.6	1357	10	US-09-815-242-11997	Sequence 11997, A
39	34	54.0	207	9	US-09-891-139A-11	Sequence 11, App1
40	34	54.0	217	9	US-10-008-118A-14	Sequence 14, App1
41	34	54.0	217	10	US-09-443-704-14	Sequence 14, App1
42	34	54.0	238	10	US-09-925-300-941	Sequence 941, App
43	34	54.0	457	9	US-10-086-135-3	Sequence 3, App1
44	34	54.0	480	9	US-09-859-888-5	Sequence 5, App1
45	34	54.0	480	9	US-09-859-888-6	Sequence 6, App1

## ALIGNMENTS

RESULT 1  
US-09-738-626-5787  
; Sequence 5787, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 5787  
; TYPE: PRT  
; LENGTH: 316  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-5787

Query Match 61.9%; Score 39; DB 9; Length 316;  
Best Local Similarity 77.8%; Pred. No. 12;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 LDENGLEFAP 12  
DB 167 VDENGKLEFAP 175

```
RESULT 2
US-09-903-456-40
; Sequence 40, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-09-903-456-40
```

```
Query Match          58.7%; Score 37; DB 10; Length 174;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 0; Gaps 0;
```

```
OY 1 IPVLDENGLF 10
    :|:| ||||
Db 80 VPIARNGLF 89
```

```
RESULT 3
US-09-903-456-26
; Sequence 26, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-09-903-456-26
```

```
Query Match          58.7%; Score 37; DB 10; Length 178;
Best Local Similarity 60.0%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 0; Gaps 0;
```

```
OY 1 IPVLDENGLF 10
    :|:| ||||
Db 3 VPIARNGLF 12
```

```
RESULT 4
US-09-903-456-38
; Sequence 38, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-09-903-456-38
```

```
Query Match          58.7%; Score 37; DB 10; Length 219;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 0; Gaps 0;
```

```
OY 1 IPVLDENGLF 10
    :|:| ||||
Db 10 VPIARNGLF 19
```

```
RESULT 5
US-09-903-456-28
; Sequence 28, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Mortierella alpina
; NAME/KEY: VARIANT
; LOCATION: (280)...(280)
; OTHER INFORMATION: Xaa = Unknown or other at position 280
US-09-903-456-28
```

```
Query Match          58.7%; Score 37; DB 10; Length 280;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 0; Gaps 0;
```

OY 1 IPVLDENGLF 10  
:1:1 ||||  
Db 70 VPIIARNGLF 79

RESULT 6

US-09-903-456-59  
; Sequence 59, Application US/09903456  
; Patent No. US20020138874A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Mukerji, Pradip  
; APPLICANT: Leonard, Amanda Eun-Yeong  
; APPLICANT: Huang, Yung-Sheng  
; APPLICANT: Pereira, Suzette L.  
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF  
; FILE REFERENCE: 6407.US.P3  
; CURRENT APPLICATION NUMBER: US/09/903,456  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: US 09/624,670  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: US 09/379,095  
; PRIOR FILING DATE: 1999-08-23  
; PRIOR APPLICATION NUMBER: US 09/145,828  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 59  
; LENGTH: 286  
; TYPE: PRT  
; ORGANISM: Mortierella alpina  
US-09-903-456-59

Query Match 58.7%; Score 37; DB 10; Length 286;  
Best Local Similarity 60.0%; Pred. No. 23;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 IPVLDENGLF 10  
:1:1 ||||  
Db 77 VPIIARNGLF 86

RESULT 7

US-09-903-456-21  
; Sequence 21, Application US/09903456  
; Patent No. US20020138874A1

GENERAL INFORMATION:

APPLICANT: Abbott Laboratories  
APPLICANT: Mukerji, Pradip  
APPLICANT: Leonard, Amanda Eun-Yeong  
APPLICANT: Huang, Yung-Sheng  
APPLICANT: Pereira, Suzette L.  
TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF  
FILE REFERENCE: 6407.US.P3  
CURRENT APPLICATION NUMBER: US/09/903,456  
CURRENT FILING DATE: 2001-07-11  
PRIOR APPLICATION NUMBER: US 09/624,670  
PRIOR FILING DATE: 2000-07-24  
PRIOR APPLICATION NUMBER: US 09/379,095  
PRIOR FILING DATE: 1999-08-23  
PRIOR APPLICATION NUMBER: US 09/145,828  
PRIOR FILING DATE: 1998-09-02  
NUMBER OF SEQ ID NOS: 116  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 21  
LENGTH: 289  
TYPE: PRT  
ORGANISM: Mortierella alpina  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (289)...(289)  
OTHER INFORMATION: Xaa = Unknown or Other at position 289

US-09-903-456-21

Query Match 58.7%; Score 37; DB 10; Length 289;  
Best Local Similarity 60.0%; Pred. No. 24;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 IPVLDENGLF 10  
:1:1 ||||  
Db 79 VPIIARNGLF 88

RESULT 8

US-09-903-456-34  
; Sequence 34, Application US/09903456  
; Patent No. US20020138874A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Mukerji, Pradip  
; APPLICANT: Leonard, Amanda Eun-Yeong  
; APPLICANT: Huang, Yung-Sheng  
; APPLICANT: Pereira, Suzette L.  
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF  
; FILE REFERENCE: 6407.US.P3  
; CURRENT APPLICATION NUMBER: US/09/903,456  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: US 09/624,670  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: US 09/379,095  
; PRIOR FILING DATE: 1999-08-23  
; PRIOR APPLICATION NUMBER: US 09/145,828  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 34  
; LENGTH: 289  
; TYPE: PRT  
; ORGANISM: Mortierella alpina  
US-09-903-456-34

Query Match 58.7%; Score 37; DB 10; Length 289;  
Best Local Similarity 60.0%; Pred. No. 24;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 IPVLDENGLF 10  
:1:1 ||||  
Db 80 VPIIARNGLF 89

RESULT 9

US-09-903-456-36  
; Sequence 36, Application US/09903456  
; Patent No. US20020138874A1

GENERAL INFORMATION:

APPLICANT: Abbott Laboratories  
APPLICANT: Mukerji, Pradip  
APPLICANT: Leonard, Amanda Eun-Yeong  
APPLICANT: Huang, Yung-Sheng  
APPLICANT: Pereira, Suzette L.  
TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF  
FILE REFERENCE: 6407.US.P3  
CURRENT APPLICATION NUMBER: US/09/903,456  
CURRENT FILING DATE: 2001-07-11  
PRIOR APPLICATION NUMBER: US 09/624,670  
PRIOR FILING DATE: 2000-07-24  
PRIOR APPLICATION NUMBER: US 09/379,095  
PRIOR FILING DATE: 1999-08-23  
PRIOR APPLICATION NUMBER: US 09/145,828  
PRIOR FILING DATE: 1998-09-02  
NUMBER OF SEQ ID NOS: 116  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 36  
LENGTH: 291  
TYPE: PRT

```
; ORGANISM: Mortierella alpina
US-09-903-456-36

Query Match
Best Local Similarity 58.7%; Score 37; DB 10; Length 291;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IPVLDENGLF 10
:|:| ||||
Db 82 VPIIARNGLF 91

RESULT 10
US-09-903-456-19
; Sequence 19, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR FILING DATE: 1999-08-23
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Mortierella alpina
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (293)...(293)
; OTHER INFORMATION: Xaa = Unknown or Other at position 293
US-09-903-456-19

Query Match
Best Local Similarity 58.7%; Score 37; DB 10; Length 293;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

1 IPVLDENGLF 10
:|:| ||||
83 VPIIARNGLF 92

RESULT 11
US-09-903-456-33
; Sequence 33, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
```

```
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Mortierella alpina
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (301)...(301)
; OTHER INFORMATION: Xaa = Unknown or Other at position 301
US-09-903-456-33

Query Match
Best Local Similarity 58.7%; Score 37; DB 10; Length 301;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IPVLDENGLF 10
:|:| ||||
Db 91 VPIIARNGLF 100

RESULT 12
US-09-903-456-13
; Sequence 13, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-09-903-456-13

Query Match
Best Local Similarity 58.7%; Score 37; DB 10; Length 317;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IPVLDENGLF 10
:|:| ||||
Db 108 VPIIARNGLF 117

RESULT 13
US-09-903-456-25
; Sequence 25, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
```

```
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Moritella alpina
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (318)...(318)
; OTHER INFORMATION: xaa = Unknown or Other at position 318
US-09-903-456-25

Query Match          58.7%; Score 37; DB 10; Length 318;
Best Local Similarity 60.0%; Pred. No. 26;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 IPVDENG 10
       :|:| ||||
Db      108 VPILRNGLF 117

RESULT 14
US-09-990-415A-8
; Sequence 8, Application US/09990415A
; Patent No. US20020165182A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Protein Cluster I
; FILE REFERENCE: 00349
; CURRENT APPLICATION NUMBER: US/09/990,415A
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 322
; TYPE: PRT
; ORGANISM: human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (25)..(25)
; OTHER INFORMATION: Xaa=A,T,G or C
US-09-990-415A-8

Query Match          58.7%; Score 37; DB 9; Length 322;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 IPVDENG 8
       :|| ||||
Db      205 IPVTDENG 212

RESULT 15
US-09-978-295A-442
; Sequence 442, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Bolstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavian, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630P1C11
; CURRENT APPLICATION NUMBER: US/09/978,295A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
```

PRIOR APPLICATION NUMBER: 60/080165  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080194  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080327  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080328  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080333  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080334  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/081070  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081049  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081071  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081195  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081955  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081819  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081952  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081838  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082568  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082569  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082704  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082804  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082700  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082797  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082796  
 PRIOR FILING DATE: 1998-04-23  
 PRIOR APPLICATION NUMBER: 60/083336  
 PRIOR FILING DATE: 1998-04-27  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083392  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083495  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083496  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083499  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083554  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083558  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083559  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083500  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083742  
 PRIOR FILING DATE: 1998-04-30  
 PRIOR APPLICATION NUMBER: 60/084366

PRIOR FILING DATE: 1998-05-05  
 PRIOR APPLICATION NUMBER: 60/084414  
 PRIOR FILING DATE: 1998-05-06  
 PRIOR APPLICATION NUMBER: 60/084441  
 PRIOR FILING DATE: 1998-05-06  
 PRIOR APPLICATION NUMBER: 60/084637  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084639  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084640  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084598  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084627  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084643  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/085339  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085338  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085323  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085582  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085700  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085689  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085579  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085580  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085573  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085704  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085697

Query Match 58.7% Score 37; DB 9; Length 436;  
 Best Local Similarity 50.0% Pred. No. 38;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 IPVLDENGLEFAP 12  
 Db 338 IPLCDEDEYKFP 349

Search completed: January 31, 2003, 08:05:15  
 Job time : 12 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2003, 07:59:05 : Search time 55 Seconds  
(Without alignments)  
20.975 Million cell updates/sec

Title: US-09-699-224A-1  
Perfect score: 63  
Sequence: 1 IPVLDENGLEFAP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seg length: 0  
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pirl:\*  
2: pirl:\*  
3: pirl:\*  
4: pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	45	71.4	359	2	C84983
2	41	65.1	180	2	A98321
3	41	65.1	707	2	T35196
4	40	63.5	157	1	F69185
5	40	63.5	337	2	B97506
6	40	63.5	337	2	AF2724
7	39	61.9	140	2	H97350
8	39	61.9	450	2	AB3238
9	39	61.9	700	2	G75318
10	39	61.9	764	2	S64951
11	39	61.9	1621	2	A82255
12	39	61.9	2893	2	A64556
13	38	60.3	162	2	T08505
14	38	60.3	197	2	A90664
15	38	60.3	157	2	F85514
16	38	60.3	197	2	AC0929
17	38	60.3	268	2	G69623
18	38	60.3	298	2	E97298
19	38	60.3	322	1	RGBYC3
20	38	60.3	365	2	AF0864
21	38	60.3	620	2	D84618
22	38	60.3	894	2	E82221
23	37.5	59.5	1659	2	H97926
24	37.5	59.5	1659	2	G95057
25	37	58.7	144	2	F83988
26	37	58.7	324	2	AB0982
27	37	58.7	328	2	F91183
28	37	58.7	328	2	B86030
29	37	58.7	328	2	C65154

30	37	58.7	347	2	AG2251	translation initia
31	37	58.7	357	2	I55210	tricarboxylate car
32	37	58.7	365	2	A98088	membrane-bound lyl
33	37	58.7	365	2	C85933	membrane-bound lyl
34	37	58.7	365	2	A65064	membrane-bound lyl
35	37	58.7	390	2	T44324	hypothetical prote
36	37	58.7	402	2	B97420	probable acyl-CoA
37	37	58.7	402	2	AC2638	acyl-CoA thiolase
38	37	58.7	443	2	S38327	serine/threonine-s
39	37	58.7	443	2	C85140	serine/threonine-s
40	37	58.7	468	2	T19628	hypothetical prote
41	37	58.7	504	2	A23282	RAD52 protein - ye
42	37	58.7	712	2	T48961	hypothetical prote
43	37	58.7	789	2	S33056	probable infected
44	37	58.7	809	2	B87260	sensory box/GSDF
45	37	58.7	1051	2	C93367	conserved hypothet

## ALIGNMENTS

## RESULT 1

C84983  
hypothetical protein mlta [imported] - Buchnera sp. (strain APS)  
C:Species: Buchnera sp.  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
C:Accession: C84983  
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.  
Nature 407, 81-86, 2000  
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp  
A:Reference number: AB4930; MUID:20445173; PMID:10993077  
A:Accession: C84983  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-359 <STO>  
A:Cross-references: GB:AP000398; GSPDB:GN00144  
A:Experimental source: strain APS  
C:Genetics:  
A:Gene: mlta; B0458

Query Match 71.4%; Score 45; DB 2; Length 359;  
Best Local Similarity 70.0%; Pred. No. 2;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IPVLDENGLEF 10  
DB 298 IPVLDENGLEF 307

## RESULT 2

A98321  
hypothetical 15.0K protein in codo 3'region (orf6) [imported] - Agrobacterium tumefaciens  
C:Species: Agrobacterium tumefaciens  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002  
C:Accession: A98321  
R:Gocher, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz,  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium  
A:Reference number: A97359; PMID:11743194  
A:Accession: A98321  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-180 <KOR>  
A:Cross-references: GB:AE007870; PIDN:AAK90091.1; PID:915160078; GSPDB:GN00170  
C:Genetics:  
A:Gene: AGR L 3039  
A:Map position: linear chromosome

Query Match 65.1%; Score 41; DB 2; Length 180;  
Best Local Similarity 50.0%; Pred. No. 5;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IPVLDENGLFAP 12  
:|||||:|  
Db 7 VPALNENGFRYP 18

## RESULT 3

T35196  
Probable DNA gyrase chain B - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 21-Jan-2000  
C:Accession: T35196  
R:Seeger, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, April 1998  
A:Reference number: Z21571  
A:Accession: T35196  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-707 <SEE>  
A:Cross-references: EMBL:AI022374; PIDN:CAI18520.1; GSPDB:GN00070; SCOEDB:SC5B8.12  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: gyrB; SCOEDB:SC5B8.12  
C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B

Query Match 65.1%; Score 41; DB 2; Length 707;  
Best Local Similarity 58.3%; Pred. No. 24;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IPVLDENGLFAP 12  
:|||||:|  
Db 290 VPVLDENGMTP 301

## RESULT 4

F69185  
Yhcv homolog MTH644 - Methanobacterium thermoautotrophicum (strain Delta H)  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: F69185  
R:Smith, D.R.; Doucette-Stamm, L.A.; DeJonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Oluf, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A:Reference number: A69000; MUID:98037514; PMID:9371463  
A:Accession: F69185  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-157 <MTH>  
A:Cross-references: GB:AE000844; GB:AE000666; NID:g2621707; PIDN:AAB85149.1; PID:g262172  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH644  
C:Superfamily: conserved hypothetical protein yhcV; CBS homology  
F:108-156/Domain: CBS homology <CBS>

Query Match 63.5%; Score 40; DB 1; Length 157;  
Best Local Similarity 63.6%; Pred. No. 6;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IPVLDENGLFAP 11  
:|||||:|  
Db 133 LPVLDENGRLA 143

## RESULT 5

B97506  
Peptide ABC transporter, ATP-binding protein [imported] - Agrobacterium tumefaciens (str  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: B97506  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium  
A:Reference number: A97359; PMID:11743194  
A:Accession: B97506  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-337 <KUR>  
A:Cross-references: GB:AE007869; PIDN:AAK87003.1; PID:g15156245; GSPDB:GN00169  
C:Genetics:  
A:Gene: AGR C.2214  
A:Map position: circular chromosome

Query Match 63.5%; Score 40; DB 2; Length 337;  
Best Local Similarity 70.0%; Pred. No. 16;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VLDENGFLAP 12  
:|||||:|  
Db 272 MLDGTGLYAP 281

## RESULT 6

AF2724  
Hypothetical protein Atu1201 [imported] - Agrobacterium tumefaciens (strain C58, Dupo  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AF2724  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo  
erger, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayav, T.; Levy, R.; Li, M.; McCl  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AF2724  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-337 <KUR>  
A:Cross-references: GB:AE008688; PIDN:AAI42212.1; PID:g17739605; GSPDB:GN00186  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu1201  
A:Map position: circular chromosome

Query Match 63.5%; Score 40; DB 2; Length 337;  
Best Local Similarity 70.0%; Pred. No. 16;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VLDENGFLAP 12  
:|||||:|  
Db 272 MLDGTGLYAP 281

## RESULT 7

H97350  
Two CBS domain containing protein [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: H97350  
R:Molling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, O.; Gibson, R.; L  
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: H97350  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-140 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK81595.1; PID:g15026776; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC3674



C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: A64556  
R:Tombl, J.F.; White, O.; Kariyave, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalek, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney, L.; Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: A64556  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-2893 <TOM>  
A:Cross-references: GB:AE000547; GB:AE000511; NID:g2313377; PIDN:AAD07355.1; PID:g231338

Query Match 61.9%; Score 39; DB 2; Length 2893;  
Best Local Similarity 66.7%; Pred. No. 2.8e+02;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 PVLVDENGLFAP 12  
|||:||||  
DB 2077 IPNKGKGLFAP 2088

RESULT 13  
T08505  
trbH protein - Enterobacter aerogenes plasmid R751  
C:Species: Enterobacter aerogenes  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: T08505  
R:Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.  
A:Title: Conservation of the genetic switch between replication and transfer genes of In  
A:Reference number: Z16434; MUID:97118926; PMID:8954881  
A:Accession: T08505  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-162 <THO>  
A:Cross-references: EMBL:U67194; NID:g1572520; PIDN:AAC64449.1; PID:g1572554  
C:Genetics:  
A:Gene: trbH  
A:Genome: plasmid R751

Query Match 60.3%; Score 38; DB 2; Length 162;  
Best Local Similarity 63.6%; Pred. No. 16;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

2 PVLVDENGLFAP 12  
|||:||||  
144 PVLQNGTAP 154

RESULT 14  
A90664  
hypothetical protein Ecs0281 [imported] - Escherichia coli (strain O157:H7, substrain R  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: A90664  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: A90664  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-197 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA833704.1; PID:g13359738; GSPDB:GN00154  
C:Genetics:  
A:Gene: Ecs0281  
C:Superfamily: phage T4 tail fiber assembly protein gp38

Query Match 60.3%; Score 38; DB 2; Length 197;  
Best Local Similarity 87.5%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PVLVDENGL 9  
|||:||||  
DB 3 PVLVDENGL 10

RESULT 15  
F85514  
unknown protein from prophage CP-933H [imported] - Escherichia coli (strain O157:H7,  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: F85514  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
hiller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamousis, K.; Apoda  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: F85514  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-197 <STO>  
A:Cross-references: GB:AE005174; NID:g12513030; PIDN:AAG54578.1; GSPDB:GN00145; UMGCP:  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z0316  
C:Superfamily: phage T4 tail fiber assembly protein gp38

Query Match 60.3%; Score 38; DB 2; Length 197;  
Best Local Similarity 87.5%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PVLVDENGL 9  
|||:||||  
DB 3 PVLVDENGL 10

Search completed: January 31, 2003, 08:01:32  
Job time : 58 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2003, 07:59:05 ; Search time 11 Seconds

(without alignments)  
45.247 Million cell updates/sec

Title: US-09-699-224A-1

Sequence: 1 IPVLDENGLFAP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt-40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	71.4	359	1	MLTA_BUCAI
2	40	63.5	986	1	AC15_DROME
3	38	60.3	203	1	GSTA_RHLE
4	38	60.3	237	1	PURT_METAC
5	38	60.3	268	1	FLRP_BACSU
6	38	60.3	322	1	SNF4_YEAST
7	38	60.3	712	1	DNL1_RHOM
8	37	58.7	322	1	SEFX1_HUMAN
9	37	58.7	322	1	SEFX1_MOUSE
10	37	58.7	322	1	SEFX1_RAT
11	37	58.7	324	1	TKRA_ECO57
12	37	58.7	324	1	TKRA_ECOLI
13	37	58.7	365	1	MLTA_ECOLI
14	37	58.7	443	1	MHR_ARATH
15	37	58.7	490	1	LAG3_CAEEL
16	37	58.7	504	1	RA52_YEAST
17	37	58.7	789	1	PRTP_EBV
18	36	57.1	181	1	PAAD_PYRHO
19	36	57.1	200	1	SOMA_HETFO
20	36	57.1	423	1	YAXJ_RHISN
21	36	57.1	562	1	AMY2_DICTH
22	36	57.1	598	1	UVRG_CHLTR
23	36	57.1	1281	1	VLB5_CAEEL
24	35.5	56.3	200	1	COAE_STRCO
25	35	55.6	123	1	UR2_RAT
26	35	55.6	192	1	PABA_STRLI
27	35	55.6	217	1	GTWU_RABIT
28	35	55.6	226	1	RPIA_METTH
29	35	55.6	229	1	RPIA_PYRAB
30	35	55.6	229	1	RPIA_PYRHO
31	35	55.6	287	1	PHNS_DESFR
32	35	55.6	314	1	CPRA_PYRHO
33	35	55.6	321	1	GUTQ_ECOLI

34	35	55.6	419	1	MANA_PSEFL
35	35	55.6	439	1	ACMA_LACIA
36	35	55.6	441	1	YXEK_BACSU
37	35	55.6	665	1	SYGB_RICPR
38	35	55.6	691	1	DNL1_MYCTU
39	35	55.6	746	1	TDCR_ECOLI
40	35	55.6	752	1	CO2_HUMAN
41	35	55.6	752	1	NSF2_DROME
42	35	55.6	1066	1	Z295_HUMAN
43	35	55.6	1357	1	RPOB_PSEAE
44	35	55.6	1357	1	RPOB_PSEPU
45	35	55.6	1375	1	RPOB_VIBCH

## ALIGNMENTS

## RESULT 1

ID	MLTA_BUCAI	STANDARD:	PRT:	359 AA.
AC	P57531:			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DE	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Membrane-bound lytic murein transglycosylase A homolog (EC 3.2.1.-)			
DE	(Murein hydrolase A).			
GN	MLTA OR B0458.			
OS	Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum)			
OS	symbiotic bacterium).			
OC	Bacteria; Proteobacteria; gamma subdivision; Buchnera.			
OX	NCBI_TaxID=118099;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Tokyo 1998;			
RX	MEDLINE=20445173; PubMed=10993077;			
RA	Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;			
RT	Genome sequence of the endocellular bacterial symbiont of aphids			
RT	Buchnera sp. APS.;			
RL	Nature 407:81-86(2000).			

-1- FUNCTION: MUREIN-DEGRADING ENZYME. MAY PLAY A ROLE IN RECYCLING OF MUROPOLYMER DURING CELL ELONGATION AND/OR CELL DIVISION (BY SIMILARITY).

-1- CATALYTIC ACTIVITY: Cleavage of the beta-1,4-glycosidic bond between N-acetylmuramic acid and N-acetylglucosamine residues, thereby conserving the energy in a newly synthesized

1,6-anhydrobond in the muramic acid residue.  
-1- SUBCELLULAR LOCATION: IN CLOSELY RELATED BACTERIA THIS PROTEIN IS ATTACHED TO THE OUTER MEMBRANE BY A LIPID ANCHOR. THIS IS APPARENTLY NOT THE CASE HERE.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----

EMBL: AP001119; BAB13155.1; -  
DR InterPro: IPR005300; MLTA.  
DR Pfam: PF03562; MLTA; 1.  
KW Cell wall; Hydrolase; Glycosidase; Complete proteome.  
SQ SEQUENCE 359 AA; 41773 MW; 24E12A277BD351AB CR64;

Query Match 71.4%; Score 45; DB 1; Length 359;

Best local Similarity 70.0%; Pred. No. 0.66;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IPVLDENGLF 10  
DB 298 IPLVDNGVF 307

RESULT 2  
AC15\_DROME STANDARD; PRT; 986 AA.  
AC P35600; 002031;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Activator 1 140 kDa subunit (Replication factor C large subunit)  
DE (Germ-line transcription factor 1).  
GN GNF1 OR CG1119.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
NCBI\_TaxID=7227;  
[1]  
RP SEQUENCE FROM N.A., AND DNA-BINDING ACTIVITY.  
RX MEDLINE=98371221; PubMed=9705493;  
RA Allen B.L., Uhlmann F., Gaur L.K., Mulder B.A., Posey K.L.,  
RA Jones L.B., Hardin S.H.;  
RA "DNA recognition properties of the N-terminal DNA binding domain  
within the large subunit of replication factor C.";  
RA Nucleic Acids Res. 26:3877-3882(1998).  
[2]  
RP SEQUENCE FROM N.A.  
RA Frank L.H., Cohen R.S.;  
RT "Cloning and characterization of a putative transcription factor  
active during oogenesis and embryogenesis.";  
RT Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.  
[3]  
RP SEQUENCE FROM N.A.  
RA STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Ceuliker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-T., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Bencze P.V., Bereman B.P., Bhandal D., Bolshakov S.,  
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A.S., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferrara C., Ferriera S., Fleischmann W.,  
RA Fostler C., Gabrielian A.F., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svitzkas R., Tector C., Turner R., Venter A.H., Wang A.H., Wang X.,  
RA Wang Z.-Y., Massaman D.A., Weinstein G.M., Weisenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
[4]  
RP INTERACTION WITH PCNA.

RX MEDLINE=97153138; PubMed=8999859;  
RA Mossi R., Jonsson Z.O., Allen B.L., Hardin S.H., Hubscher U.;  
RT "Replication factor C interacts with the C-terminal side of  
RT proliferating cell nuclear antigen.";  
RL J. Biol. Chem. 272:1769-1776(1997).  
CC -1- FUNCTION: THE ELONGATION OF PRIMED DNA TEMPLATES BY DNA POLYMERASE  
CC DELTA AND EPSILON REQUIRES THE ACTION OF THE ACCESSORY PROTEINS  
CC PCNA AND ACTIVATOR 1. THE 140 SUBUNIT BINDS TO THE PRIMER-TEMPLATE  
CC JUNCTION (BY SIMILARITY).  
CC -1- SUBUNIT: INTERACTS WITH C-TERMINUS OF PCNA.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -1- SIMILARITY: BELONGS TO THE ACTIVATOR 1 140 KDA SUBUNIT FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: L17340; AAA28573.1; -  
DR EMBL: U97685; AAB58311.1; -  
DR EMBL: AE003604; AAF52082.1; -  
DR FlyBase: FBgn0004913; Gnf1.  
DR InterPro: IPR003959; AAA\_ATPase\_centr.  
DR InterPro: IPR001357; BRCT.  
DR InterPro: IPR000862; RFCdomain.  
DR Pfam: PF00004; AAA; 1.  
DR Pfam: PF00533; BRCT; 1.  
DR SMART: SM00292; BRCT; 1.  
DR PROSITE: PS50172; BRCT; 1.  
KW DNA replication; ATP-binding; Transcription regulation; DNA-binding;  
KW Activator; Nuclear protein; Zinc-finger.  
FT DOMAIN 232 313  
FT NP\_BIND 487 494  
FT ZN\_FING 585 602  
FT DOMAIN 955 959  
FT CONFIDE 559 559  
SQ SEQUENCE 986 AA; 108614 MW; A1E9CEB0879545F CMC64;  
Query Match 63.5%; Score 40; DB 1; Length 986;  
Best Local Similarity 70.0%; Pred. No. 17;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
DB 297 TPILSEDLF 306  
OY 1 TPVLDENGLP 10  
DB 297 TPILSEDLF 306  
RESULT 3  
GSTA\_RHILE  
ID GSTA\_RHILE STANDARD; PRT; 203 AA.  
AC Q52828;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE GSTA protein.  
OS Rhizobium leguminosarum.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=384;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B155;  
RA Al-Kafaf N.K., Tawfiq N.K.;  
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; X89816; CAA61942.1; -  
DR HSSP; P30712; ILJR.  
DR InterPro: IPR004046; GST\_Cterm.  
DR InterPro: IPR004045; GST\_Nterm.  
DR Pfam; PF00043; GST\_C; 1.  
DR Pfam; PF02798; GST\_N; 1.  
DR Transferase.  
SQ SEQUENCE 203 AA; 21987 MW; 7A4F3D94E6B282E1 CRC64;

Query Match 60.3%; Score 38; DB 1; Length 203;  
Best Local Similarity 75.0%; Pred. No. 7.5;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 IPVDENG 8  
:|||||  
Db 52 VPVDDNG 59

RESULT 4  
PUR7\_METAC STANDARD; PRT: 237 AA.

AC Q8TIS; 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Phosphoribosylaminimidazole-succinocarboxamide synthase (EC 6.3.2.6)  
DE (SAICAR synthetase).  
GN PURC OR MA4063.  
OS Methanosarcina acetivorans.  
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;  
OC Methanosarcinaceae; Methanosarcina.

OX NCBI\_Taxid=2214;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=C2A / ATCC 35395 / DSM 2834;  
RX MEDLINE=21929760; PubMed=11932238;

RA Galagan J.E., Nishbaum C., Roy A., Endritz M.G., Macdonald P.,  
RA Fitzhugh W., Galvo S., Engels R., Smirnov S., Atroch D., Brown A.,  
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,  
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,  
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,  
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,  
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,  
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,  
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,  
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
RA Metcalf W.W., Britten B.;

RT "The genome of Methanosarcina acetivorans reveals extensive metabolic  
RT and physiological diversity.";  
RL Genome Res. 12:532-542(2002).

CC -1- CATALYTIC ACTIVITY: ATP + 1-(5-phosphoribosyl)-4-carboxy-5-  
CC aminimidazole + L-aspartate = ADP + phosphate + 1-(5-  
CC phosphoribosyl)-4-(N-succinocarboxamide)-5-aminimidazole.  
CC -1- PATHWAY: De novo purine biosynthesis; seventh step.

CC -1- SIMILARITY: BELONGS TO THE SAICAR SYNTHETASE FAMILY.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE011118; AM07411.1; ALT\_INT.  
DR PROSITE; PS01057; SAICAR\_SYNTHETASE\_1; FALSE\_NEG.  
DR PROSITE; PS01058; SAICAR\_SYNTHETASE\_2; 1.

KW Purine biosynthesis; Ligase; Complete proteome.  
SQ SEQUENCE 237 AA; 27092 MW; 56CC6698B747AA22 CRC64;

Query Match 60.3%; Score 38; DB 1; Length 237;  
Best Local Similarity 50.0%; Pred. No. 8.8;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

1 IPVDENGFLRP 12  
:|||||  
Db 162 VPVDEKGLLP 173

RESULT 5  
FLRP\_BACSU STANDARD; PRT: 268 AA.

AC P39753; P94582; 01-FEB-1995 (Rel. 31, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Flagellar hook-basal body complex protein flrp.  
GN FLRP.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI\_Taxid=1423;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98015417; PubMed=9353933;  
RA Prescan E., Moszer I., Boursier L., Cruz Ramos H.C., De La Fuente V.,  
RA Hullo M.-F., Leiong C., Schleich S., Sekowska A., Song B.H.,  
RA Villani G., Kunst F., Danchin A., Glaser P.;

RT "The Bacillus subtilis genome from gercb (311 degrees) to lhc (334  
RT degrees).";  
RL Microbiology 143:3313-3328(1997).

RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogatawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Boriss R., Boursier L., Brans A., Braun M., Britnell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
RA Ertlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N.,  
RA Ghim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,  
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Potwollig S., Prescott A.M.,  
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale F.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,  
RA Sekiuchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumschein E., Yoshikawa H., Danchin A.;

RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
RT subtilis.";  
RL Nature 390:249-256(1997).

RN [3]  
RP SEQUENCE OF 1-135 FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=95138040; PubMed=7836311;

```

RA Abhayawardhane Y., Stewart G.C.;
RT "Bacillus subtilis possesses a second determinant with extensive
sequence similarity to the Escherichia coli mreB morphogene.";
RL J. Bacteriol. 177:765-773(1995).
CC
CC -1- SIMILARITY: BELONGS TO THE FLAGELLA BASAL BODY ROD PROTEINS
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z83337; CAB05941.1; -
DR EMBL; Z99122; CAB15656.1; -
DR EMBL; U12962; AAA67880.1; -
DR Subtilist; BG10918; flhp.
DR InterPro: IPR001444; flag_db-rod.
DR Pfam: PF00460; flg_db-rod; 1.
DR PROSITE: PS00588; FLAGELLA_BB_ROD; 1.
KW Flagella; Complete proteome. LAMP -> GRLG (IN REF. 3).
FT CONFLICT 68 71
SQ SEQUENCE 268 AA; 29400 MW; 13C0CB83D0CB4AA3 CRC64;

Query Match 60.3%; Score 38; DB 1; Length 268;
Best local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PVIDENG 8
Db 142 PVIDENG 148

RESULT 6
SNF4_YEAST STANDARD; PRT; 322 AA.
ID SNF4_YEAST
AC P12904;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nuclear protein SNF4 (Regulatory protein CAT3).
GN SNF4 OR CAT3 OR YGL15W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
RA MEDLINE=89006284; PubMed=3049255;
RA Schueller H.-J., Ertian K.-D.;
RT "Molecular characterization of yeast regulatory gene CAT3 necessary
RT for glucose derepression and nuclear localization of its product.";
RL Gene 67:247-257(1988).
[2]
SEQUENCE FROM N.A.
RA MEDLINE=90097921; PubMed=2481228;
RA Celenza J.L., Eng F.J., Carlson M.;
RT "Molecular analysis of the SNF4 gene of Saccharomyces cerevisiae:
RT evidence for physical association of the SNF4 protein with the SNF1
RT protein kinase.";
RL Mol. Cell. Biol. 9:5045-5054(1989).
[3]
SEQUENCE FROM N.A.
RA Lauguin G.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE OF 1-21 FROM N.A.
RA Doi A., Doi K.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE OF 30-34 AND 316-322.

```

```

RX MEDLINE=94131988; PubMed=7905477;
RA Mitchellhill K.I., Stapleton D., Gao G., House C., Mitchell B.,
RA Katsis F., Witters L.A., Kemp B.E.;
RT "Mammalian AMP-activated protein kinase shares structural and
RT functional homology with the catalytic domain of yeast Snf1 protein
RT kinase.";
RL J. Biol. Chem. 269:2361-2364(1994).
CC
CC -1- FUNCTION: THIS PROTEIN CAUSES EXPRESSION OF GLUCOSE-REPRESSIBLE
CC GENES UPON GLUCOSE DEPRIVATION. IT INTERACTS AND HAS FUNCTIONAL
CC RELATIONSHIP TO THE PROTEIN-KINASE SNF1.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA
CC SUBUNIT FAMILY.
CC -1- SIMILARITY: CONTAINS 4 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21760; AAA34472.1; -
DR EMBL; M30470; AAA35061.1; -
DR EMBL; Z72637; CAA96823.1; -
DR EMBL; D16506; BAA03958.1; -
DR PIR; JTO316; RCBYCS.
DR PIR; A33480; A33480.
DR SGD; S0003083; SNF4.
DR InterPro: IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Carbohydrate metabolism; Transcription regulation; Nuclear protein;
KW Repeat; CBS domain.
FT DOMAIN 35 89 CBS 1.
FT DOMAIN 117 175 CBS 2.
FT DOMAIN 192 246 CBS 3.
FT DOMAIN 259 318 CBS 4.
SQ SEQUENCE 322 AA; 36401 MW; 51B387E346EE9561 CRC64;

Query Match 60.3%; Score 38; DB 1; Length 322;
Best local Similarity 62.5%; Pred. No. 12;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 IPVVDENG 8
Db 223 VPVDENG 230

RESULT 7
DNLJ_RHOMR STANDARD; PRT; 712 AA.
ID DNLJ_RHOMR
AC P49421;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA ligase (EC 6.5.1.2) (Polydeoxyribonucleotide synthase [NAD+]).
GN LIGA OR LIG.
OS Rhodothermus marinus (Rhodothermus obamensis).
OC Bacteria; Bacteroidetes; Sphingobacteriales;
OC Crenotrichaceae; Rhodothermus.
OX NCBI_TaxID=29549;
[1]
SEQUENCE FROM N.A.
RA STRAIN=R-21;
RC MEDLINE=95369716; PubMed=7642120;
RA Thorbjarnardottir S.H., Jonsson Z.O., Andresson O.S.,
RA Kristjansson J.K., Eggertsson G., Paladottir A.;
RT "Cloning and sequence analysis of the DNA ligase-encoding gene of
RT Rhodothermus marinus, and overproduction, purification and
RT characterization of two thermophilic DNA ligases.";
RL Gene 161:1-6(1995).

```



CC -1- FUNCTION: THIS PROTEIN CATALYZES THE FORMATION OF PHOSPHODIESTER  
CC LINKAGES BETWEEN 5'-PHOSPHORYL AND 3'-HYDROXYL GROUPS IN DOUBLE-  
CC STRANDED DNA USING NAD AS A COENZYME AND AS THE ENERGY SOURCE FOR  
CC THE REACTION. IT IS ESSENTIAL FOR DNA REPLICATION AND REPAIR OF  
CC DAMAGED DNA. THIS ENZYME IS THERMOSTABLE BEING ACTIVE AT 5-75  
CC DEGREES CELSIUS WITH APPARENT OPTIMAL ACTIVITY ABOVE 55 DEGREES  
CC CELSIUS.  
CC -1- CATALYTIC ACTIVITY: NAD(+) + (deoxyribonucleotide)(n) +  
CC (deoxyribonucleotide)(m) = AMP + nicotinamide nucleotide +  
CC (deoxyribonucleotide)(n+m).  
CC -1- MISCELLANEOUS: THIS ENZYME IS THERMOSTABLE.  
CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT DNA LIGASE FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U10483; AAA93198.1; -.  
CC HSSP: O87703; 1B04.  
CC InterPro: IPR001357; BRCT.  
CC InterPro: IPR004150; DNA\_ligase\_OB.  
CC InterPro: IPR001679; DNAligase.  
CC InterPro: IPR003265; Endo\_3c.  
CC InterPro: IPR003583; HNH\_1.  
CC InterPro: IPR000445; HNH.  
CC InterPro: IPR004149; znf\_DNAligase\_C4.  
CC Pfam: PF00533; BRCT; 1.  
CC Pfam: PF00633; HNH; 1.  
CC Pfam: PF01653; DNA\_ligase\_N; 1.  
CC Pfam: PF03119; DNA\_ligase\_ZBD; 1.  
CC Pfam: PF03120; DNA\_ligase\_OB; 1.  
CC ProDom: PD003944; DNAligase; 1.  
CC SMART: SM00292; BRCT; 1.  
CC SMART: SM00478; ENDO3c; 1.  
CC SMART: SM00278; HNH; 3.  
CC SMART: SM00332; LIGANC; 1.  
CC TIGRFAMS: TIGR00575; dnlj; 1.  
CC PROSITE: PS50172; BRCT; 1.  
CC PROSITE: PS01055; DNA\_LIGASE\_N1; 1.  
CC PROSITE: PS01056; DNA\_LIGASE\_N2; 1.  
CC LIGASE, DNA repair; DNA replication; NAD.  
CC FT DOMAIN 633 712 BRCT.  
CC FT BINDING 153 153  
CC FT SEQUENCE 712 AA; 79487 MW; 893AD3A78F7FEFC1 CRC64;  
CC  
CC Query Match 60.3%; Score 38; DB 1; Length 712;  
CC Best Local Similarity 77.8%; Pred. No. 28;  
CC Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IPVDENGL 9  
|||:|||||  
Db 697 IPMDDEGL 705

RESULT 8  
SF\_X1\_HUMAN STANDARD: PRT; 322 AA.  
ID SF\_X1\_HUMAN STANDARD: PRT; 322 AA.  
AC O9H9B4; O9HA53.  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Sideroflexin 1.  
GN SFXN1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Li N., Chen T., Wan T., Zhang W., Cao X.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland, and Teratocarcinoma;  
RA Nishigaki T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishigaki T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Watanabe S., Hosokawa M., Hosokawa T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakami K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,  
RA Niinomiya K., Iwayanagi T.;  
RT "NDO human cDNA sequencing project."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 19-322 FROM N.A.  
RC TISSUE=Lung;  
RA Strausberg R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Might be involved in the transport of a component  
CC required for iron utilization into or out of the mitochondria.  
CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE SIDEROFLEXIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF327346; BAU56007.1; -.  
CC EMBL: AK022287; BAB14003.1; ALT\_SEQ.  
CC EMBL: AK022938; BAB14318.1; ALT\_INIT.  
CC EMBL: BC020517; AAH20517.1; ALT\_INIT.  
CC InterPro: IPR004686; MTC.  
CC TIGRFAMS: TIGR00798; mtc; 1.  
CC KW Transport; Iron transport; Iron; Mitochondrion; Transmembrane.  
CC FT TRANSMEM 103 120 POTENTIAL.  
CC FT TRANSMEM 147 167 POTENTIAL.  
CC FT TRANSMEM 175 195 POTENTIAL.  
CC FT TRANSMEM 229 249 POTENTIAL.  
CC FT TRANSMEM 267 287 POTENTIAL.  
CC FT TRANSMEM 267 287 POTENTIAL.  
CC FT CONFLICT 1 63 MSGELPPINKKEPRWDQSTFGRANHFVTVDPRIILLN  
CC FT CONFLICT 1 63 EQLESARKIVHDYDQGIIVPOL -> MI (IN REF. 2;  
CC FT CONFLICT 1 63 BAB14003).  
CC FT CONFLICT 73 73 Y-> K (IN REF. 2; BAB14318).  
CC FT CONFLICT 291 322 SMSGVSTLEAFLOAKIDESHLELRVFNKGL -> RYFV  
CC FT CONFLICT 291 322 ITRONHEEY (IN REF. 2; BAB14003).  
CC FT SEQUENCE 322 AA; 35619 MW; 47E0317FE27990DB CRC64;  
CC  
CC Query Match 58.7%; Score 37; DB 1; Length 322;  
CC Best Local Similarity 87.5%; Pred. No. 19;  
CC Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IPVDENG 8  
||| |||||  
Db 205 IPVTDENG 212

RESULT 9  
SF\_X1\_MOUSE STANDARD: PRT; 322 AA.  
ID SF\_X1\_MOUSE STANDARD: PRT; 322 AA.  
AC O99JRI; O9CZG4;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Sideroflexin 1.  
GN SFXN1 OR F.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DISEASE.  
 RX MEDLINE=21172735; PubMed=11274051;  
 RA Fleming M.D., Campagna D.R., Haslett J.N., Trenor C.C. III,  
 RA Andrews N.C.;  
 RT "A mutation in a mitochondrial transmembrane protein is responsible  
 RT for the pleiotropic hematological and skeletal phenotype of  
 RT flexed-tail (f/f) mice."  
 RL Genes Dev. 15:652-657(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Araiawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barsh G.,  
 RA Blake J., Botelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast;  
 RA Strausberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Might be involved in the transport of a component  
 CC required for iron utilization into or out of the mitochondria.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -1- TISSUE SPECIFICITY: Widely expressed, with highest expression in  
 CC kidney and liver.  
 CC -1- DEVELOPMENTAL STAGE: Very high levels in the liver during the  
 CC period of embryonic hepatic hemopoiesis.  
 CC -1- DISEASE: Defects in SFXN1 are the cause of a transitory  
 CC hypochromic, microcytic anemia characterized by a large number of  
 CC siderocytes containing non-heme iron granules. The anemia begins  
 CC at 12 dpc, is most intense at 15 dpc and is still severe at birth,  
 CC but disappears by 2 weeks of age. Mutant adults are no longer  
 CC anemic, but they have an impaired response to hemopoietic stress.  
 CC Most homozygotes also have flexed tails and a belly spot.  
 CC -1- SIMILARITY: BELONGS TO THE SIDEROFLEXIN FAMILY.  
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A  
 CC FRAMESHIFT IN POSITION 282.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AF325260; AAK39428.1; -  
 CC EMBL: AK012650; BAB28382.1; ALT\_FRAME.  
 CC EMBL: BC005743; AAH05743.1; -  
 CC MGD: MGI:2137677; Sfxn1.  
 CC InterPro: IPR004686; Mtc.  
 CC TrIRFAMS: TIGR00798; mtc; 1.  
 CC Transport: Iron transport; Iron; Mitochondrion; Transmembrane.

FT TRANSMEM 103 120 POTENTIAL.  
 FT TRANSMEM 147 167 POTENTIAL.  
 FT TRANSMEM 175 195 POTENTIAL.  
 FT TRANSMEM 229 249 POTENTIAL.  
 FT TRANSMEM 267 287 POTENTIAL.  
 SO SEQUENCE 322 AA; 35649 MW; E3B055CB803CDEFA7 CRC64;  
 Query Match 58.7%; Score 37; DB 1; Length 322;  
 Best Local Similarity 87.5%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 IPVDENG 8  
 DB 205 IPVDENG 212  
 RESULT 10  
 ID SFXN1\_RAT STANDARD; PRT; 322 AA.  
 AC 063965;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Sideroflexin 1 (Tricarboxylate carrier protein).  
 GN SFXN1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=94179133; PubMed=8132491;  
 RA Azzi A., Glerum M., Koller R., Mertens W., Spycher S.;  
 RT "The mitochondrial tricarboxylate carrier."  
 RL J. Bioenerg. Biomembr. 25:515-524(1993).  
 CC -1- FUNCTION: Might be involved in the transport of a component  
 CC required for iron utilization into or out of the mitochondria.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -1- TISSUE SPECIFICITY: Expressed in liver and brain.  
 CC -1- SIMILARITY: BELONGS TO THE SIDEROFLEXIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: S70011; AAB30258.1; ALT\_INT.  
 CC InterPro: IPR004686; Mtc.  
 CC TrIRFAMS: TIGR00798; mtc; 1.  
 CC Transport: Iron transport; Iron; Mitochondrion; Transmembrane.  
 FT TRANSMEM 103 120 POTENTIAL.  
 FT TRANSMEM 147 167 POTENTIAL.  
 FT TRANSMEM 175 195 POTENTIAL.  
 FT TRANSMEM 229 249 POTENTIAL.  
 FT TRANSMEM 267 287 POTENTIAL.  
 SO SEQUENCE 322 AA; 35546 MW; D28CBDB88EAB5CD CRC64;  
 Query Match 58.7%; Score 37; DB 1; Length 322;  
 Best Local Similarity 87.5%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 IPVDENG 8  
 DB 205 IPVDENG 212  
 RESULT 11  
 ID TKRA\_EC057 STANDARD; PRT; 324 AA.  
 ID TKRA\_EC057

AC P58220;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 2-ketoglucuronate reductase (EC 1.1.1.215) (2KR) (2-ketoadonate reductase).  
 GN TKRA OR Z4978 OR ECS4438.  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;  
 RL "Genome sequence of enterohemorrhagic Escherichia coli O157:H7." Nature 409:529-533(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 050952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Ikeda T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T., Kuhara S., Shiba T., Hattori M., Shingawa H.;  
 RL "Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12." DNA Res. 8:11-22(2001).  
 CC -1- FUNCTION: CATALYZES THE NADPH-DEPENDENT REDUCTION OF 2,5-DIKETO-D-GLUCONATE (25DKG) TO 5-KETO-D-GLUCONATE (5KDG), 2-KETO-D-GLUCONATE (2KDG) TO D-GLUCONATE (GA), AND 2-KETO-L-GLUCONATE (2KLG) TO L-IDONATE (IA) (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: D-gluconate + NADP(+) = 2-dehydro-D-gluconate + NADPH.  
 CC -1- SUBUNIT: HOMODIMER (PROBABLE).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID DEHYDROGENASES FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AE005582; AAG58702.1; ALT\_INIT.  
 CC EMBL: AP002565; BAB37861.1; ALT\_INIT.  
 CC HSSP: P36234; IGDH.  
 CC InterPro: IPR002162; D\_2hyd-dh.  
 CC pfam: PF02826; 2-Hacid-DH; 1.  
 CC pfam: PF02826; 2-Hacid-DH; 1.  
 CC PROSITE: PS00065; D\_2-HYDROXYACID-DH\_1; FALSE\_NEG.  
 CC PROSITE: PS00670; D\_2-HYDROXYACID-DH\_2; 1.  
 CC PROSITE: PS00671; D\_2-HYDROXYACID-DH\_3; 1.  
 CC Oxidoreductase: Gluconate utilization; NADP: Complete proteome.  
 FT ACT\_SITE 237 SUBSTRATE-BINDING (BY SIMILARITY).  
 FT ACT\_SITE 266 BY SIMILARITY.  
 FT ACT\_SITE 285 BY SIMILARITY.  
 SQ SEQUENCE 324 AA; 35396 MW; A6EB339119DD250 CRC64;

Query Match 58.7%; Score 37; DB 1; Length 324;  
 Best local Similarity 70.0%; Pred. No. 19;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PVIDENGLFA 11

Db 239 PVIDENGLFA 248  
 11:11111  
 RESULT 12  
 ID TKRA\_ECOLI  
 AC P37666;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 2-ketoglucuronate reductase (EC 1.1.1.215) (2KR) (2-ketoadonate reductase).  
 GN TKRA OR B3553.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=94316500; PubMed=8041620;  
 RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;  
 RL "Analysis of the Escherichia coli genome. V. DNA sequence of the region from 76.0 to 81.5 minutes." Nucleic Acids Res. 22:2576-2586(1994).  
 RN [2]  
 RP REVISIONS TO C-TERMINUS.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;  
 RL "The complete genome sequence of Escherichia coli K-12." Science 277:1453-1474(1997).  
 RN [3]  
 RP SEQUENCE OF 1-22, AND CHARACTERIZATION.  
 RC STRAIN=K12 / W3110;  
 RX MEDLINE=99030322; PubMed=9811658;  
 RA Tum D.-Y., Lee B.-Y., Hamm D.-H., Pan J.-G.;  
 RL "The ylaE gene, located at 80.1 minutes on the Escherichia coli chromosome, encodes a 2-ketoadonate reductase." J. Bacteriol. 180:5984-5988(1998).  
 RN [4]  
 RP PARTIAL SEQUENCE OF 1-8.  
 RC STRAIN=K12 / EMG2;  
 RX MEDLINE=97443975; PubMed=9298646;  
 RA Link A.J., Robison K., Church G.M.;  
 RL "Comparing the predicted and observed properties of proteins encoded in the genome of Escherichia coli K-12." Electrophoresis 18:1259-1313(1997).  
 CC -1- FUNCTION: CATALYZES THE NADPH-DEPENDENT REDUCTION OF 2,5-DIKETO-D-GLUCONATE (25DKG) TO 5-KETO-D-GLUCONATE (5KDG), 2-KETO-D-GLUCONATE (2KDG) TO D-GLUCONATE (GA), AND 2-KETO-L-GLUCONATE (2KLG) TO L-IDONATE (IA).  
 CC -1- CATALYTIC ACTIVITY: D-gluconate + NADP(+) = 2-dehydro-D-gluconate + NADPH.  
 CC -1- SUBUNIT: HOMODIMER (PROBABLE).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -1- MISCELLANEOUS: OPTIMAL PH IS 7.5.  
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID DEHYDROGENASES FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: U00039; AAB18530.1; ALT\_FRAME.

DR EMBL: AE00432; AAC76577.1; ALT\_INIT.  
DR HSSP: P36234; 1GDH.  
DR Ecocore: EGI2272; tkra.  
DR InterPro: IPR002162; D\_2hyddec\_dh.  
DR Pfam: PF02826; 2-Hacid\_dh; 1.  
DR PROSITE: PS00065; D\_2-HYDROXYACID\_DH\_1; FALSE\_NEG.  
DR PROSITE: PS00670; D\_2-HYDROXYACID\_DH\_2; 1.  
DR PROSITE: PS00671; D\_2-HYDROXYACID\_DH\_3; 1.  
KM Oxidoreductase; Gluconate utilization; NADP: Complete proteome.  
FT ACT\_SITE 237 237  
FT ACT\_SITE 266 266 BY SIMILARITY.  
FT ACT\_SITE 285 285 BY SIMILARITY.  
SQ SEQUENCE 324 AA; 35395 MM; 1B21339B1337D255 CRC64;  
  
Query Match 58.7%; Score 37; DB 1; Length 324;  
Best local Similarity 70.0%; Pred. No. 19;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
2 PVLDENGLFA 11  
11:111111  
239 PVVDENALIA 248  
  
RESULT 13  
MLTA\_ECOLI STANDARD; PRT; 365 AA.  
AC P46885; P76638; Q46928;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Membrane-bound lytic murein transglycosylase A precursor (EC 3.2.1.-)  
GN (Murein hydrolase A) (MLT38).  
DE MLTA OR MLT OR B2813 OR Z4130 OR ECS3673.  
OS Escherichia coli, and  
OC Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562, 83334;  
[1]  
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RC STRAIN-K12;  
RA Lomatzsch J., Tempin M.F., Kraft A.R., Vollmer W., Hoeltje J.-V.;  
RT "Outer membrane localization of murein hydrolases: MLTA, a third  
lipoprotein lytic transglycosylase in Escherichia coli.";  
RL J. Bacteriol. 179:5465-5470(1997).  
[2]  
SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mao B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
[3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;  
RA MEDLINE-2107495; PubMed-1120651;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blatner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
RL Nature 409:529-533(2001).  
[4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-O157:H7 / RIMD 0509952;  
RA MEDLINE-21156231; PubMed-11258796;

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tope T.,  
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
RA Kuhara S., Shiba T., Hattori M., Shingawa H.,  
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22(2001).  
[5]  
RP SEQUENCE OF 147-161; 204-213 AND 258-280, AND CHARACTERIZATION.  
RA MEDLINE-94117367; PubMed-8288527;  
RA Ursinus A., Hoeltje J.-V.;  
RT "Purification and properties of a membrane-bound lytic  
transglycosylase from Escherichia coli.";  
RL J. Bacteriol. 176:338-343(1994).  
[6]  
RP INTERACTION WITH MIPA AND MRGB/PONB.  
RC STRAIN-ATCC 53338 / MC1061;  
RX MEDLINE-99156961; PubMed-10037771;  
RA Vollmer W., von Rechenberg M., Hoeltje J.-V.;  
RT "Demonstration of molecular interactions between the murein polymerase  
PBPB, the lytic transglycosylase MLTA, and the scaffolding protein  
MIPA of Escherichia coli.";  
RL J. Biol. Chem. 274:6726-6734(1999).  
CC -1- FUNCTION: MUREIN-DEGRADING ENZYME. MAY PLAY A ROLE IN RECYCLING OF  
CC MUROPOLYMER DURING CELL ELONGATION AND/OR CELL DIVISION. OPTIMAL  
CC ACTIVITY IS BETWEEN PH 4.0 AND 4.5; LOSES ITS ACTIVITY RAPIDLY AT  
CC TEMPERATURES ABOVE 30 DEGREES CELSIUS. DEGRADES MUREIN GLYCAN  
CC STRANDS AND INSOLUBLE, HIGH-MOLECULAR WEIGHT MUREIN SACCULI.  
CC -1- CATALYTIC ACTIVITY: Cleavage of the beta-1,4-glycosidic bond  
CC between N-acetylmuramic acid and N-acetylglucosamine residues,  
CC thereby conserving the energy in a newly synthesized  
CC 1,6-anhydrobond in the muramic acid residue.  
CC -1- SUBUNIT: FORMS A TRIMERIC COMPLEX WITH MRGB/PONB AND MIPA IN  
CC VITRO.  
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid  
CC anchor.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
DR EMBL: U32224; AAC45723.1; -;  
DR EMBL: AE000365; AAC75855.1; -;  
DR EMBL: U29581; AAB40463.1; ALT\_INIT.  
DR EMBL: AE005509; AAG57927.1; -;  
DR EMBL: AP002563; BAB37096.1; -;  
DR Ecocore: EGI3085; MLTA.  
DR InterPro: IPR005300; MLTA.  
DR Pfam: PF03562; MLTA; 1.  
DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
KW Cell wall; Hydrolase; Glycosidase; Signal; Lipoprotein;  
KW Outer membrane; Multigene family; Complete proteome.  
FT SIGNAL 1 20  
FT CHAIN 21 365  
FT LIPID 21 21  
FT CONFICT 346 346 H -> N (IN REF. 1).  
SQ SEQUENCE 365 AA; 40410 MM; 5ECBB92C1B8D5969 CRC64;  
  
Query Match 58.7%; Score 37; DB 1; Length 365;  
Best local Similarity 60.0%; Pred. No. 21;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 PVLDENGLFA 10  
11:111111  
Db 295 PVLDNNGKF 304  
  
RESULT 14

MRK\_ARATH STANDARD: PRT: 443 AA.  
ID MRK\_ARATH  
AC P43294: 09SV67; 09C5D4;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Serine/threonine-protein kinase MRK (EC 2.7.1.1).  
GN MRK OR ATG13020 OR P25G13.110.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OC NCBI\_Taxid=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=94032493; PubMed=8218420;  
RA Moran T.V., Walker J.C.;  
RT "Molecular cloning of two novel protein kinase genes from Arabidopsis thaliana";  
RT Biochim. Biophys. Acta 1216:9-14(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=20083488; PubMed=10617198;  
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,  
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,  
RA Harris B., Ansoorge W., Brandt P., Griwell L.A., Rieger M.,  
RA Weischelgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,  
RA Kreis B., Delsenay M., Puigdomenech P., Watson M., Schmidtheini T.,  
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
RA Vos P., Hohenseil J., Zimmermann W., Medler H., Ridley P.,  
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,  
RA van der Schueren J., Grymoprez B., Chuang Y.-J., Vandebussche F.,  
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,  
RA Weltensgeer T., Bothe G., Rampsberger U., Hilbert H., Braun M.,  
RA Holzner E., Brandt A., Peters S., van Staveren M., Dirxse W.,  
RA Moollman P., Klein lankhorst R., Rose M., Hauf J., Koelter P.,  
RA Berner S., Hempel S., Feldpausch M., Lambirth S., van den Daele H.,  
RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,  
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,  
RA Clark L., Doggett J., Hall S., Kay M., Leonard N., Mcclay K., Mayes R.,  
RA Pettitt A., Rajadream M.A., Lyne M., Benes V., Reichmann S.,  
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,  
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fartmann B., Grandert K., Dauner D., Herzi A.,  
RA Neumann S., Agiriotou A., Vitale D., Liguori R., Piravandi E.,  
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Aubourg S.,  
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
RA Chefkor F., Cooke R., Berger C., Monfort A., Casacuberta E.,  
RA Glibons T., Weber R., Vandenbol M., Barges M., Terol J., Torres A.,  
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,  
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Blatke C.,  
RA Frisman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,  
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,  
RA Stoneking T., Kalcick J., Graves T., Harmon G., Edwards J.,  
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
RA Minx P., Bentley D., Fulton L., Miller N., Greco T., Kemp K.,  
RA Krimer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,  
RA Nelson J., Spleth J., Ryan E., Andrews S., Geisel C., Layman D.,  
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,  
RA Antonou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,  
RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Matero A., Shah R.,  
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,  
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodi M., Johnson A.,  
RA Chen E., Maria M., Martienssen R., McCombie W.R.;  
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana";  
RT Nature 402:769-777(1999).  
RN [3]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RC STRAIN=cv. Columbia;  
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;  
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the  
RT SSP consortium (Saik/Standford/PGEC)";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF PLANT GROWTH AND  
CC DEVELOPMENT.  
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a  
CC short form; are produced by alternative splicing.  
CC -1- TISSUE SPECIFICITY: ROOTS, LEAVES AND STEMS.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC CDC2/CDCX SUBFAMILY, STRONGEST, TO MAMMALIAN MKK.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: L07249; AAA1854.1; -;  
DR EMBL: AL079349; CAB45501.1; -;  
DR EMBL: AL161535; CAB78344.1; -;  
DR EMBL: AF360324; AAK26034.1; -;  
DR EMBL: AY056332; AAL07181.1; -;  
DR HSSP: P2941; 1B38.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002290; Ser\_thr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR Prodom: PD000001; Euk.pkinase; 1.  
DR SMART: SM00220; S\_TKc\_1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
KW Transferrase; Serine/threonine-protein kinase; ATP-binding;  
KW Alternative splicing.  
FT DOMAIN 12 291  
FT NP\_BIND 18 26 PROTEIN KINASE.  
FT BINDING 41 41 ATP (BY SIMILARITY).  
FT ACT\_SITE 133 133 ATP (BY SIMILARITY).  
FT VARSPLIC 1 10 BY SIMILARITY.  
FT CONFLICT 66 66 MMVEFVENC -> ME (IN SHORT ISOFORM).  
FT N -> K (IN REF. 1)  
SQ SEQUENCE 443 AA; 50895 MW; AF92A2233B673B0F CRC64;  
QY 5 DENGFLAP 12  
DB 382 DENGFLAP 389  
ID LAG3\_CAEEL STANDARD: PRT: 490 AA.  
AC 009260; 09NGS2;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE LAG-3 protein;  
GN LAG-3 OR C32A3.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_Taxid=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=20289813; PubMed=10830967;  
RX Petcherski A.G., Kimble J.;  
RT "LAG-3 is a putative transcriptional activator in the C. elegans Notch  
RT pathway";

```

RL Nature 405:364-368(2000).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Thomas K.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
CC -i FUNCTION: GIP-1 and lin-12 promote signalling by recruiting lag-3
CC to target promoters, where it functions as a transcriptional
CC activator.
CC -i SUBUNIT: Associates with lag-1.
CC -i SUBCELLULAR LOCATION: Nuclear.
CC -i ALTERNATIVE PRODUCTS: 2 isoforms; a (shown here) and b; may be
CC produced by alternative splicing.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF241847; AAF71523.1; -.
DR EMBL: AF241846; AAF71522.1; -.
DR EMBL: 248241; CA88284.1; -.
DR EMBL: 248241; CAC42265.1; -.
DR WormPep: C32A3.1a; CE27810.
DR WormPep: C32A3.1b; CE01505.
KM Transcription regulation; Nuclear protein; Alternative splicing.
FT DOMAIN 162 467 GLN-RICH.
FT DOMAIN 185 188 POLY-GLN.
FT DOMAIN 341 346 POLY-GLN.
FT DOMAIN 353 357 POLY-GLN.
FT DOMAIN 372 380 POLY-GLN.
FT DOMAIN 390 400 POLY-GLN.
FT DOMAIN 433 442 POLY-GLN.
FT VARSPLIC 1 22 MISSING (IN ISOFORM B).
SQ SEQUENCE 490 AA; 56483 MW; 396A6C194C382A07 CRC64;

Query Match 58.7%; Score 37; DB 1; Length 490;
Best Local Similarity 54.5%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PVIDENGLFAP 12
Db 260 PVIDENNLAVP 270

```

ch completed: January 31, 2003, 08:00:07  
time : 13 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2003, 07:59:05 : Search time 29 Seconds  
(without alignments)  
85.261 Million cell updates/sec

Title: US-09-699-224A-1

Perfect score: 63

Sequence: 1 IPIVDENGLFAP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp rodent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	73.0	332	17	Q96ZS5 sulfolobus
2	43	68.3	136	16	Q8XIG7 Q8XIG7 clostridium
3	41	65.1	180	16	Q8UAS1 Q8UAS1 agrobacteri
4	41	65.1	245	2	Q31394 Q31394 bacillus me
5	41	65.1	707	16	Q69998 Q69998 streptomyc
6	41	65.1	2462	16	Q8RGZ3 Q8RGZ3 fusobacteri
7	41	65.1	2806	16	Q8RI19 Q8RI19 fusobacteri
8	40	63.5	157	17	Q26740 Q26740 methanobact
9	40	63.5	224	5	Q9V8G6 Q9V8G6 drosophila
10	40	63.5	337	16	Q8UG41 Q8UG41 agrobacteri
11	40	63.5	2205	5	Q9NGO2 Q9NGO2 dictyostell
12	39	61.9	103	5	Q9N7L0 Q9N7L0 leishmania
13	39	61.9	140	16	Q97D09 Q97D09 clostridium
14	39	61.9	221	5	Q9GPI8 Q9GPI8 anopheles g
15	39	61.9	224	5	Q9GPI9 Q9GPI9 anopheles g
16	39	61.9	224	5	Q8WQML Q8WQML anopheles g

17	39	61.9	225	5	Q9V8I7	Q9V8I7 drosophila
18	39	61.9	442	13	Q9PVJ9	Q9PVJ9 gallus gall
19	39	61.9	450	16	Q8U651	Q8U651 agrobacteri
20	39	61.9	569	13	Q9YHW0	Q9YHW0 gallus gall
21	39	61.9	676	5	Q9NDT5	Q9NDT5 balanus amp
22	39	61.9	700	16	Q9RS05	Q9RS05 delinococcus
23	39	61.9	764	39	Q12500	Q12500 saccharomyc
24	39	61.9	994	10	Q8SR35	Q8SR35 oryza sativ
25	39	61.9	1621	16	Q9KTA5	Q9KTA5 vibrio chol
26	39	61.9	2893	16	Q25063	Q25063 helicobacte
27	39	60.3	118	14	Q991Z7	Q991Z7 uncultured
28	39	60.3	162	2	P71183	P71183 enterobacte
29	39	60.3	162	2	Q9AHH1	Q9AHH1 commonas t
30	39	60.3	197	16	Q8Z342	Q8Z342 salmonella
31	39	60.3	197	16	Q8X7M0	Q8X7M0 escherichia
32	39	60.3	256	2	Q52745	Q52745 rumiococcu
33	39	60.3	262	17	Q8TIS9	Q8TIS9 methanosarc
34	39	60.3	278	17	Q9HLM0	Q9HLM0 thermoplasm
35	39	60.3	298	16	Q97E74	Q97E74 clostridium
36	39	60.3	333	4	Q8TCE7	Q8TCE7 homo sapien
37	39	60.3	365	16	Q8ZMC0	Q8ZMC0 salmonella
38	39	60.3	365	16	Q8Z423	Q8Z423 salmonella
39	39	60.3	407	2	Q8RM59	Q8RM59 bacteroides
40	39	60.3	565	2	P96170	P96170 vibrio para
41	39	60.3	605	3	Q9HGP1	Q9HGP1 schizosacch
42	39	60.3	620	10	Q81008	Q81008 arabidopsis
43	39	60.3	724	16	Q984Y8	Q984Y8 rhizobium l
44	39	60.3	894	16	Q9KSJ8	Q9KSJ8 vibrio chol
45	39	60.3	915	16	Q8XOR6	Q8XOR6 raistonia s

#### ALIGNMENTS

##### RESULT 1

Q96ZS5 PRELIMINARY; PRT; 332 AA.  
AC Q96ZS5;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE Hypothetical protein ST1760.  
OS Sulfolobus tokodaii.  
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
OC Sulfolobus.  
OX NCBI\_TaxID=111955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JCM 10545 / 7;  
RX PubMed=11572479;  
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
RA Sekine M., Baba S.-I., Anai A., Kosugi H., Hosoyama A., Fukui S.,  
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,  
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,  
RA Oshima T., Kikuchi H.;  
RT "Complete genome sequence of an aerobic thermophilic  
Crenarchaeon, Sulfolobus tokodaii strain7.";  
RL DNA Res. 8:123-140(2001).  
DR EMBL: AP000987; BAB56848.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 332 AA; 38801 MW; ACADIFG4C3AFEA9 CRC64;

Query Match 73.0%; Score 46; DB 17; Length 332;

Best Local Similarity 66.7%; Pred. No. 3.4;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPIVDENGLFAP 12

Db 161 IPIVDENGLIMP 172

```

RESULT 2
ID 08XIG7 PRELIMINARY; PRT; 136 AA.
AC 08XIG7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable inosine-5'-monophosphate dehydrogenase.
GN G6AB OR CPE2153.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohnani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RA "Complete genome sequence of Clostridium perfringens, an anaerobic
RA flesh-eater.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL: AP003193; BAB81859.1; -.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS_2.
DR SMART: SM00116; CBS; 2.
DR KMW Complete proteome.
SQ SEQUENCE 136 AA; 15608 MW; 36646441544E0AE CRC64;

Query Match 68.3%; Score 43; DB 16; Length 136;
Best Local Similarity 60.0%; Pred. No. 4.7;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 IPVLDENGFLP 10
Db 106 VPVVDNGIF 115

RESULT 3
ID 08UAS1 PRELIMINARY; PRT; 180 AA.
AC 08UAS1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein Atu3299.
GN ATU3299 OR AGR L3039.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RA "The genome of the natural genetic engineer Agrobacterium tumefaciens
RA C58.";
RA Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorlo B., Goldman B.S., Cao Y., Askenazi M., Halling L.,
RA Houmel K., Gordon J., Yaudin W., Tatchouk O., Epp A., Liu F.,

```

```

RA Wollam C., Allinger M., Doughly D., Scott C., Lappas C., Markelz B.,
RA Planagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RA "Genome sequence of the plant pathogen and biotechnology agent
RA RT Agrobacterium tumefaciens C58.";
RA Science 294:2323-2328(2001).
RL EMBL: AE009259; AAL44115.1; ALT_INIT.
DR EMBL: AE008353; AAK90091.1; -.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 180 AA; 19735 MW; 71D0FF440AE034F0 CRC64;

Query Match 65.1%; Score 41; DB 16; Length 180;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IPVLDENGFLP 12
Db 7 VPALNENGFTRP 18

RESULT 4
ID 031394 PRELIMINARY; PRT; 245 AA.
AC 031394;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Sensory transduction histidine kinase.
GN ORF2.
OS Bacillus megaterium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PV447;
RA Vazquez G.U., Pettinari M.J., Mendez B.M.;
RA Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL: Y14588; CAA74928.1; -.
DR InterPro: IPR000014; PAS_domain.
DR Pfam: PR00989; PAS; 1.
DR SMART: SM00091; PAS; 1.
DR TIGRPFAM: TIGR00229; sensory_box; 1.
KM Kinase.
SQ SEQUENCE 245 AA; 27391 MW; 6B6E2A39BB25E7C1 CRC64;

Query Match 65.1%; Score 41; DB 2; Length 245;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IPVLDENG 8
Db 199 IPVLDENG 206

RESULT 5
ID 069998 PRELIMINARY; PRT; 707 AA.
AC 069998;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to many e.g. GYRB-STRAU p20832 staphylococcus aureus. dna
DE gyrase subunit B.
GN SC05822 OR SC58.12.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K., Harris D.;

```



RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE-97000351; PubMed-8843436;  
 RA Redenbach M., Kieser H.M., Denapite D., Eichner A., Cullum J.,  
 RA Knaus H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Lake L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 CC -1- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING  
 OF DOUBLE-STRANDED DNA.  
 CC -1- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.  
 DR EMBL: AL022374; CAI18520.1; -.  
 DR HSSP: P06982; 1A06.  
 DR InterPro: IPR003594; ATPbind\_ATPase.  
 DR InterPro: IPR002936; DNAPrim\_toprim.  
 DR InterPro: IPR002288; DNA\_gyraseB\_C.  
 DR InterPro: IPR001241; DNA\_topoisomII.  
 DR InterPro: IPR004359; HIS\_KIN\_sig.  
 DR Pfam: PF00204; DNA\_gyraseB\_C.1.  
 DR Pfam: PF00986; DNA\_gyraseB\_C.1.  
 DR Pfam: PF02518; HATPase\_C.1.  
 DR Pfam: PF01751; Toprim; 1.  
 DR PRINTS: PR00418; TP12FAMILY.  
 DR PRODOM: PD149633; DNA\_gyraseB\_C.1.  
 DR SMART: SM00387; HATPase\_C.1.  
 DR SMART: SM00433; TOP2C.1.  
 DR PROSITE: PS00177; TOPOISOMERASE\_II; UNKNOWN\_1.  
 DR ATP-binding; Isomerase; Topoisomerase.  
 SO SEQUENCE 707 AA; 77269 MW; 1A5427EB8EFB661 CRC64;  
 Query Match 65.1%; Score 41; DB 16; Length 707;  
 Best Local Similarity 58.3%; Pred. No. 63;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Oy 1 IPVLDENGLEFAP 12  
 Db 290 VPVLDENGQMT 301  
 RESULT 6  
 Q8RGZ3  
 ID 08RGZ3 PRELIMINARY; PRT; 2462 AA.  
 AC 08RGZ3;  
 DT 01-JUN-2002 (TREMblrel. 21, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DE Hemolysin.  
 GN FN0132.  
 OS Fusobacterium nucleatum (subsp. nucleatum).  
 OC Bacteria; Fusobacteriia; Fusobacterium.  
 NX NCBI\_TaxID=76856;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 25586;  
 RX MEDLINE-21886394; PubMed-11889109;  
 RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,  
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,  
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,  
 RA Larsen N., D'Souza M., Malunas T., Pusch G., Haselkorn R.,  
 RA Fongstein M., Kyriades N., Overbeek R.;  
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium  
 nucleatum strain ATCC 25586.";  
 RL J. Bacteriol. 184:2005-2018(2002).  
 DR EMBL: AE010527; AAL94338.1; -.  
 KW Complete proteome.  
 SO SEQUENCE 2462 AA; 267525 MW; 0DE3BEFAEA5A26 CRC64;  
 Query Match 65.1%; Score 41; DB 16; Length 2462;  
 Best Local Similarity 87.5%; Pred. No. 2,4e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 IPVLDENG 8  
 Db 945 VPVLDENG 952  
 RESULT 7  
 Q8RI19  
 ID 08RI19 PRELIMINARY; PRT; 2806 AA.  
 AC 08RI19;  
 DT 01-JUN-2002 (TREMblrel. 21, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DE Hemolysin.  
 GN FN1817.  
 OS Fusobacterium nucleatum (subsp. nucleatum).  
 OC Bacteria; Fusobacteriia; Fusobacterium.  
 NX NCBI\_TaxID=76856;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 25586;  
 RX MEDLINE-21886394; PubMed-11889109;  
 RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,  
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,  
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,  
 RA Larsen N., D'Souza M., Malunas T., Pusch G., Haselkorn R.,  
 RA Fongstein M., Kyriades N., Overbeek R.;  
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium  
 nucleatum strain ATCC 25586.";  
 RL J. Bacteriol. 184:2005-2018(2002).  
 DR EMBL: AE010484; AAL93916.1; -.  
 KW Complete proteome.  
 SO SEQUENCE 2806 AA; 304991 MW; 314255277612B236 CRC64;  
 Query Match 65.1%; Score 41; DB 16; Length 2806;  
 Best Local Similarity 87.5%; Pred. No. 2,8e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 IPVLDENG 8  
 Db 1128 VPVLDENG 1135  
 RESULT 8  
 Q26740  
 ID 026740 PRELIMINARY; PRT; 157 AA.  
 AC 026740;  
 DT 01-JAN-1998 (TREMblrel. 05, Created)  
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
 DE Conserved protein.  
 GN MTH644.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteriia; Methanobacteriales;

OC Methanobacteriaceae; Methanothermobacter.  
 ON NCBI\_TaxID=187420;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,  
 RA Aldrege T., Bashirzadeh R., Blakey D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
 RA Spadifora R., Vycare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT delta: functional analysis and comparative genomics.";  
 RL J. Bacteriol. 179:7135-7155(1997).  
 DR EMBL: AE000844; AAB85149.1; -;  
 DR HSSP: P50099; 1ZFI  
 DR InterPro: IPR000644; CBS\_domain.  
 DR Pfam: PF00571; CBS; 2.  
 SMART: SM00116; CBS; 2.  
 NC Complete proteome.  
 SQ SEQUENCE 157 AA; 17387 MW; 862166DC65B43E9A CRC64;  
 Query Match 63.5%; Score 40; DB 17; Length 157;  
 Best Local Similarity 63.6%; Pred. No. 19;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 IPVLDNGLFA 11  
 Db 133 LPVIDNGRLA 143  
 RESULT 9  
 ID 09V8G6 PRELIMINARY; PRT; 224 AA.  
 AC 09V8G6;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE CG5164 protein (Glutathione S-transferase) (CH146544).  
 GN GST3 OR CG5164  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NC NCBI\_TaxID=7227;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN-BERKELEY;  
 MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Abmayri A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballow R.M., Basu A., Bayendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottaker P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Chewley S., Dahlke C., Daveport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieri S., Fleischmann W.,  
 RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Gload A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobary C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier L.E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20237668; PubMed=10773456;  
 RA Singh M., Silva E., Schultze S., Sinclair D.A.R., Fitzpatrick K.A.,  
 RA Honda B.M.;  
 RT "Cloning and characterization of a new theta-class glutathione-S-  
 RT transferase (GST) gene, gst-3, from Drosophila melanogaster.";  
 RL Gene 247:167-173(2000).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Abmayri A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nuno J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE003799; AAF57701.1; -;  
 DR EMBL: AF179869; AAF64647.1; -;  
 DR EMBL: AY058983; AAL13612.1; -;  
 DR HSSP: P30712; ILUR.  
 DR FlyBase: FBgn0034335; Gst3.  
 DR InterPro: IPR004046; GST\_Cterm.  
 DR InterPro: IPR004045; GST\_Nterm.  
 DR Pfam: PF00043; GST\_C; 1.  
 DR Pfam: PF02798; GST\_N; 1.  
 KW Transferase.  
 SQ SEQUENCE 224 AA; 24959 MW; 148C2F359DBBC806 CRC64;  
 Query Match 63.5%; Score 40; DB 5; Length 224;  
 Best Local Similarity 60.0%; Pred. No. 28;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 IPVLDNGLFA 10  
 Db 57 VPMLDNGTFF 66  
 RESULT 10  
 ID 08UG41 PRELIMINARY; PRT; 337 AA.  
 AC 08UG41;  
 DT 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE ABC transporter, nucleotide binding/ATPase protein.  
 GN ATUJ201 OR AGR\_C.2214.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 NC NCBI\_TaxID=176299;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608550; PubMed=11743193;  
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,  
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D., Sr.,

RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 RA Kuttyavin T., Levy R., Li M.-Y., McClelland E., Palmieri A.,  
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon Kamm B., Liao S., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.;  
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*  
 RT C58".  
 RL Science 294:2317-2323(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608551; PubMed=11743194;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Gurrillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Hounmel K., Gordon J., Vaudin M., Jarichouk O., Epp A., Liu F.,  
 RA Wollman C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,  
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,  
 RA Cielo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 RT *Agrobacterium tumefaciens* C58".  
 RL Science 294:2323-2328(2001).  
 DR EMBL: AE009083; AAL42212.1; -;  
 DR EMBL: AE008048; AAK87003.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 337 AA; 37353 MW; 0E9FF6FC462FCABC CRC64;  
 QY Query Match 63.5%; Score 40; DB 16; Length 337;  
 Best Local Similarity 70.0%; Pred. No. 44;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 272 MLDGTGLXAP 281  
 QY 3 VLDENGLEAP 12  
 Db 272 MLDGTGLXAP 281  
 RESULT 11  
 Q9NGO2 PRELIMINARY; PRT; 2205 AA.  
 AC Q9NGO2;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Kinesin Unc104/Kif1a homolog.  
 GN UNCI04.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20014990; PubMed=10545495;  
 RA Pollock N., de Hostos E.L., Turck C.W., Vale R.D.;  
 RT "Reconstitution of membrane transport powered by a novel dimeric  
 RT kinesin motor of the Unc104/Kif1a family purified from  
 RT Dictyostelium".  
 RL J. Cell Biol. 147:493-506(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Pollock N., Vale R.D.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF245277; AAF63384.1; -;  
 DR HSSP: P17119; 3KAR.  
 DR InterPro: IPR001753; kinesin\_motor.  
 DR InterPro: IPR001849; PH.  
 DR Pfam: PF00225; kinesin\_1.  
 DR PRINTS: PR00380; KINESINHEAVY.  
 DR SMART: SM00129; KISC. 1.  
 DR SMART: SM00233; PH. 1.  
 DR PROSITE: PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
 DR PROSITE: PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 DR PROSITE: PS50003; PH\_DOMAIN1; 1.  
 KW ATP-binding; Coiled coil; Microtubules; Motor protein.

SQ SEQUENCE 2205 AA; 248002 MW; 02C5101E9D61C9ED CRC64;  
 QY Query Match 63.5%; Score 40; DB 5; Length 2205;  
 Best Local Similarity 75.0%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Db 841 VPLIDENG 848  
 QY 1 IPVLDENG 8  
 Db 841 VPLIDENG 848  
 RESULT 12  
 Q9N7L0 PRELIMINARY; PRT; 103 AA.  
 AC Q9N7L0;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE Probable DNA ligase (Fragment).  
 GN LM28.201.  
 OS Leishmania major.  
 OC Eukaryota; Euzoenzoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OX NCBI\_TaxID=5664;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FRIEDLIN;  
 RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL390935; CAC00927.1; -;  
 DR InterPro: IPR001357; BRCT.  
 DR Pfam: PF00533; BRCT. 1.  
 DR SMART: SM00292; BRCT. 1.  
 DR PROSITE: PS50172; BRCT. 1.  
 KW Ligase.  
 FT NON\_TER 1  
 FT NON\_TER 103  
 SQ SEQUENCE 103 AA; 10513 MW; F011768639791624 CRC64;  
 QY Query Match 61.9%; Score 39; DB 5; Length 103;  
 Best Local Similarity 63.6%; Pred. No. 19;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 Db 86 VPVLDENGMA 96  
 QY 1 IPVLDENGMA 11  
 Db 86 VPVLDENGMA 96  
 RESULT 13  
 Q97D09 PRELIMINARY; PRT; 140 AA.  
 AC Q97D09;  
 DT 01-OCT-2001 (TREMBLrel. 18, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Two CBS domain containing protein.  
 GN CAC3674.  
 OS Clostridium acetobutylicum.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 OC Clostridiales; Clostridiaceae; Clostridium.  
 OX NCBI\_TaxID=1488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
 RX MEDLINE=21359325; PubMed=11466286;  
 RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,  
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitt J., Wolf Y.I.,  
 RA Tatunov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
 RA Bennett G.N., Koonin E.V., Smith D.R.;  
 RT "Genome sequence and comparative analysis of the solvent-producing  
 RT bacterium *Clostridium acetobutylicum*".  
 RL J. Bacteriol. 183:4823-4838(2001).  
 DR EMBL: AE007862; AAK81595.1; -;  
 DR InterPro: IPR000644; CBS\_domain.

DR Pfam: PF00571; CBS; 2.  
 DR SMART: SM00116; CBS; 2.  
 KW Complete proteome.  
 SO SEQUENCE 140 AA; 16349 MW; FC6964F9D375AF06 CRC64;

Query Match 61.9%; Score 39; DB 16; Length 140;  
 Best Local Similarity 60.0%; Pred. No. 26;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 IPVLDENG 10  
 |||:|:|:  
 Db 37 IPVLDNGKY 46

## RESULT 14

O9GPI8 PRELIMINARY; PRT; 221 AA.

AC O9GPI8; 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

OS Anopheles gambiae (African malaria mosquito).  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;  
 OC Anophelinae.

OX NCBI\_TaxID=7165;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=ZAN/U;

RA Ortelli F., Jensen B., Rossiter L., Wang X., Roth C.,  
 Collins F.H., Hemingway J.;

RT "Identification of a novel class of insect glutathione S-transferases  
 RT associated with DDT resistance in the malaria vector Anopheles

RT gambiae.";  
 Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF316636; AAG45164.1; -.  
 DR HSSP: P46088; 2GSO.

DR InterPro: IPR004046; GST\_Cterm.  
 DR InterPro: IPR004045; GST\_Nterm.

DR Pfam: PF000043; GST\_C; 1.  
 DR Pfam: PF02798; GST\_N; 1.  
 DR Transferase.

SO SEQUENCE 221 AA; 24850 MW; ABC124C784079FC5 CRC64;

Query Match 61.9%; Score 39; DB 5; Length 221;  
 Best Local Similarity 87.5%; Pred. No. 43;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 IPVLDENG 8  
 |||:|:|:  
 Db 55 IPVLDNG 62

## RESULT 15

O9GPI9 PRELIMINARY; PRT; 224 AA.

AC O9GPI9; 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

OS Anopheles gambiae (African malaria mosquito).  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;  
 OC Anophelinae.

OX NCBI\_TaxID=7165;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=ZAN/U;

RA Ortelli F., Ranson H., Jensen B., Rossiter L., Wang X., Roth C.,  
 Collins F.H., Hemingway J.;

RT "Identification of a novel class of insect glutathione S-transferases

RT associated with DDT resistance in the malaria vector Anopheles  
 RT gambiae.";  
 Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF316635; AAG45163.1; -.  
 DR HSSP: P30712; 1LJR.

DR InterPro: IPR004046; GST\_Cterm.  
 DR InterPro: IPR004045; GST\_Nterm.

DR Pfam: PF000043; GST\_C; 1.  
 DR Pfam: PF02798; GST\_N; 1.  
 DR Transferase.

SO SEQUENCE 224 AA; 25284 MW; B7D3047BC2ACB1A6 CRC64;

Query Match 61.9%; Score 39; DB 5; Length 224;  
 Best Local Similarity 87.5%; Pred. No. 43;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IPVLDENG 8  
 |||:|:|:  
 Db 55 IPVLDNG 62

Search completed: January 31, 2003, 08:02:08  
 Job time : 32 secs